

# Impacto del Programa de Alimentación Escolar en el Trabajo Infantil

*Una aproximación desde la toma de decisiones familiares*<sup>\*</sup>

Indira M. Porto Gutiérrez

Directores: Jorge A. Gallego y José A. Guerra

## Resumen

Esta investigación analiza el impacto del Programa de Alimentación Escolar en el trabajo infantil en Colombia a través de varias técnicas de evaluación de impacto que incluyen emparejamiento simple, emparejamiento genético y emparejamiento con reducción de sesgo. En particular, se encuentra que este programa disminuye la probabilidad de que los escolares trabajen alrededor de un 4 %. Además, se explora que el trabajo infantil se reduce gracias a que el programa aumenta la seguridad alimentaria, lo que consecuentemente cambia las decisiones de los hogares y anula la carga laboral en los infantes. Son numerosos los avances en primera infancia llevados a cabo por el Estado, sin embargo, estos resultados sirven de base para construir un marco conceptual en el que se deben rescatar y promover las políticas públicas alimentarias en toda la edad escolar.

---

<sup>\*</sup>Documento presentado como tesis para optar por el título de Maestría en Economía para la Universidad del Rosario. email: [porto.indira@urosario.edu.co](mailto:porto.indira@urosario.edu.co)

# 1. Introducción

El Programa de Alimentación Escolar (PAE) es uno de los programas sociales más antiguos de Colombia. Fue creado en 1926 en el departamento de Boyacá, extendiéndose luego a nivel nacional en los años treinta. El programa está dirigido a todos los colegios públicos, en primaria y secundaria así como en el área rural y urbana y tiene una cobertura aproximada de 4 millones de niños y jóvenes. El PAE consiste en servir refrigerios, desayunos o almuerzos, en promedio 120 días de los 180 del calendario escolar. De igual manera, se priorizan las comunidades rurales, población desplazada e indígena y colegios con mayor proporción de escolares en los niveles 1 y 2 del Sisbén<sup>1</sup> (Ley 1176 de 2007). El objetivo del PAE es contribuir a mejorar el desempeño académico de las niñas, niños y adolescentes, logrando su asistencia regular, y promoviendo la formación de hábitos alimenticios saludables en la población escolar, con la participación activa de la familia, la comunidad y el Estado por medio de los entes territoriales (Lineamientos técnicos PAE, 2007).

En ese sentido el PAE se convierte en una política que resulta llamativa a la hora de investigar sobre su impacto, especialmente si se tiene en cuenta que sus objetivos se encuentran dirigidos a mejorar varias esferas del entorno infantil (escolar, familiar, nutricional, etc). Aunque en el programa se invierten recursos cercanos al billón de pesos anual (0.19 % del PIB), la cobertura en Colombia llega tan solo al 52 %, por debajo de países de la región como Brasil y Cuba, donde la cobertura es cercana al 100 % (III Seminario Latinoamericano de Alimentación Escolar, 2010).

Poniendo en contexto la variable de interés de esta investigación, en Colombia según la Gran Encuesta Integrada de Hogares, participaron en el mercado laboral 1.149.300 niños y adolescentes en 2014 (9,7 %) y se espera reducir esa cifra a 7,9 % en el 2018. Lo anterior es clave en un contexto donde la reducción del trabajo infantil se convierte en el blanco de política de entidades internacionales como la Organización Internacional del Trabajo (OIT) y el Fondo de las Naciones Unidas para la Infancia, que han atado la reducción en trabajo infantil con una mayor asistencia y mejor desempeño escolar (UNICEF)<sup>2</sup>. Dichas recomendaciones se han trasladado al contexto nacional y son puntos importantes de los últimos planes de desarrollo que coinciden en resaltar

---

<sup>1</sup>El nivel 1 y 2 del Sisbén corresponde a la población más pobre y vulnerable del país.

<sup>2</sup>Por ejemplo, la OIT en el 2012 desarrolló la Política Pública para la prevención y erradicación de Trabajo Infantil en Colombia. Mientras que la UNICEF, en el 2011 publicó el documento "Trabajo Infantil y Juvenil: de la formulación de los derechos a su aplicación"

que el trabajo infantil aleja a las niñas y niños de la escuela y afecta las condiciones que permiten un buen desempeño escolar (PND 2010-2014, PND 2014-2018).

Siguiendo esa idea, la pregunta de investigación es cuál es el impacto del Programa de Alimentación Escolar sobre el trabajo infantil y la hipótesis a evaluar es que el PAE disminuye el trabajo infantil. El mecanismo es que *el programa reduce el costo de oportunidad de mantenerse en el colegio*: el niño no tiene que ir a trabajar para conseguir el alimento diario debido a que el PAE se lo proporciona. Esto a su vez cambia las decisiones y el comportamiento de los hogares, haciendo que sea menos rentable enviar a los niños a trabajar. En particular, bajo una aproximación de seguridad alimentaria se demostrará cómo el programa cuando disminuye el costo de estudiar de los niños, hace que las familias se sienten más seguras, capaces de enviar a los niños al colegio y alejarlos del mercado laboral. Dicho razonamiento se comprobará mostrando que el programa aumenta la seguridad alimentaria de los hogares, cambio que se ve complementado con un correspondiente aumento de la asistencia escolar y el nivel de actividad física de los niños y las niñas.

Como se analizará más adelante la contribución de esta investigación además de analizar el impacto de un programa sobre variables de resultados no convencionales, también es realizar un aporte a la literatura, en la medida en que trata de unir dos enfoques académicos en una misma idea: aquel que entiende la educación de los hijos como un costo de oportunidad y aquel que justifica las inversiones e intervenciones en edades tempranas como el PAE. Así mismo, esta investigación es novedosa a la hora de entender el componente multidimensional del programa y alejarse de las evaluaciones de impacto convencionales que analizan variables tradicionales de tipo nutricional y físico.

Los datos provienen de la Encuesta Nacional de Situación Nutricional para el año 2010. Esta encuesta es representativa a nivel nacional y tiene información de características socio demográficas de los niños y de sus padres. Para esta investigación solo se utiliza la submuestra de los niños y jóvenes entre 5 y 17 años que pertenecen a colegios públicos<sup>3</sup>.

En cuanto a la estrategia empírica, en primer lugar se usará un modelo de Mínimos Cuadrados Ordinarios en el que se utilizarán variables referenciadas a lo largo de la literatura (Atanasio et al, 2004; Quiroga, 2006). A su vez, ya que el propósito de este estudio es estimar una aproximación causal del PAE, se sigue un recuento de metodologías cuasi-experimentales que van desde

---

<sup>3</sup>Por el diseño de la encuesta la submuestra sigue siendo representativa a nivel nacional (ENDS, 2010)

emparejamiento simple hasta emparejamiento con reducción de sesgo y emparejamiento genético, métodos con los que se encuentra un coeficiente robusto del impacto del programa.

Precisamente, en esta investigación se encuentran resultados robustos de que el PAE reduce el trabajo infantil cerca de un 4 %, utilizando diferentes tipos de emparejamiento y controles. Esto, además de ser consistente con los objetivos principales del programa, también podría determinar una nueva agenda en la que se promueva de manera más intensiva las herramientas que brinda el PAE.

Este artículo se compone de 7 secciones incluyendo esta introducción, la segunda parte relaciona literatura sobre programas de alimentación escolar y efectos en los niños, además brinda una explicación de algunos conceptos claves para esta investigación. En la tercera sección se expone el marco teórico con el que se interpretarán los resultados. En la cuarta sección se presentan la base de datos y las estadísticas descriptivas. En la quinta se plantea y se detalla la estrategia empírica y en la sexta se presentan el análisis de los resultados. Finalmente, en la séptima sección se presentan las conclusiones.

## **2. Conceptos y Trabajos Previos**

### **2.1. Conceptos**

Lo primero en este trabajo es traer a colación algunas definiciones útiles, específicamente explicar en qué consiste el Programa de Alimentación Escolar (PAE) y qué es el Trabajo Infantil. En primer lugar, el PAE consiste en el suministro de un complemento nutricional con alimentos para los niños, niñas y adolescentes matriculados en el sistema educativo público. El servicio prestado por el PAE se realiza durante el calendario escolar, y son las instituciones educativas los escenarios donde se presta el servicio de alimentación en forma directa a los estudiantes.

A su vez, el objetivo general del PAE, es el desarrollo de un conjunto de acciones alimentarias, nutricionales, de salud y de formación en adecuados hábitos alimenticios que buscan mejorar las condiciones de los escolares y apoyar su vinculación y permanencia en el sistema educativo. A nivel de objetivos específicos el programa busca garantizar el mínimo del 20 % de las reco-

mendaciones diarias de nutrientes, así como fomentar la asistencia regular de los niños, niñas y adolescentes, disminuir la deserción escolar y contribuir a mejorar el aprendizaje de los infantes y jóvenes, reduciendo el hambre. Finalmente, el programa también busca promover y fomentar la formación de hábitos alimenticios que favorezcan la salud en todas las etapas del ciclo de la vida, mediante procesos formativos en el ámbito escolar.

La población objetivo a la cual se dirige el PAE son los niños, niñas y adolescentes de las áreas rurales y urbanas. Sin embargo, ante la magnitud de escolares inscritos en el Sistema Público de Educación se aplica una estrategia de focalización para la selección de los beneficiarios del programa. Este proceso consta de dos etapas: en la primera se escogen las instituciones educativas y en la segunda los niños participantes.

La primera etapa se realiza en los Consejos Municipales de Política Social, en el Comité de Infancia y Familia, o en el de Seguridad Alimentaria y Nutricional. En esta selección se priorizan las instituciones educativas que prestan el nivel de preescolar, que imparten educación primaria a grupos étnicos, con alta proporción de población desplazada, o ambos. Luego se prioriza las instituciones en el área rural y los colegios urbanos con alta concentración de población en niveles 1 y 2 de Sisbén.

Por otro lado, la segunda etapa corresponde a la selección de beneficiarios dentro de la institución educativa. Se debe priorizar los niños que cursan preescolar, primaria, que habitan en el área rural y a población indígena. Así mismo, a los estudiantes en situación de desplazamiento sin importar los grados en que estén matriculados (estos deben estar cubiertos al 100%). Adicionalmente, las instituciones educativas urbanas asignan el complemento cubriendo la totalidad de niños, niñas y adolescentes matriculados en cada grado, pero dando prelación a los matriculados en grados inferiores y con necesidades especiales, hasta llegar a los superiores. En resumen, se seleccionan de manera prioritaria los niños y niñas en preescolar, primero y segundo de primaria, y los pertenecientes a pueblos indígenas y en situación de desplazamiento. Luego, en las regiones en donde los anteriores grupos prioritarios estén cubiertos, se podrán atender en su orden, la cobertura de los grados tercero, cuarto y quinto de primaria y sexto en adelante (Lineamientos técnicos PAE, 2007).

Por su parte, la definición relevante de Trabajo Infantil para esta investigación se toma de la Organización Internacional del Trabajo (OIT). Esta señala que el término trabajo infantil es defi-

nido como aquel que priva a los niños de su infancia, su potencial y su dignidad, y que es nocivo para su desarrollo físico y mental. Se refiere a la labor que siendo perjudicial para el niño, interfiere en su escolarización y le niega la oportunidad de ir a la escuela, obligándolo a abandonar prematuramente las aulas, o exigiendo que intente combinar la asistencia a la escuela con jornadas de trabajo (OIT, 2004). Por lo tanto, se considera trabajo infantil a todas esas actividades que obstaculizan el adecuado desarrollo de los niños y jóvenes, que de alguna manera les resten esfuerzos para participar en actividades propias de su edad como ir al colegio, hacer las tareas, divertirse con sus amigos y desarrollar completamente su capacidad a través de las relaciones interpersonales y familiares. En particular, para Colombia el Ministerio de Trabajo en 2011 publicó el Informe sobre el comportamiento del Trabajo Infantil colombiano, cuya metodología consideraba a un niño o joven como ocupado si trabaja por lo menos una hora a la semana, reciba o no remuneración<sup>4</sup>.

Ahora bien, diversas áreas han desarrollado estudios que establecen que las inversiones durante la niñez pueden llegar a ser un factor determinante del desempeño de las personas a lo largo de su vida. Se destaca que el resultado de dichas intervenciones se manifiesta tanto en el corto como en el largo plazo; un programa de alimentación escolar, por ejemplo, aumenta la nutrición de los niños en el corto plazo, mientras que los efectos de esta intervención a nivel cognitivo y social se manifiesten gradualmente en el largo plazo. Estos efectos resumen algunos resultados expuestos en la literatura, aquellos de los descubrimientos más relevantes se mostrarán a continuación, así como hechos estilizados relacionados con la importancia de la infancia y cómo esta determina muchos aspectos futuros de los agentes.

## **2.2. Revisión de literatura**

En esta parte se muestra la importancia de las inversiones en etapas tempranas; luego como eje central de esta investigación, se explora la idea de la educación como una decisión de inversión de

---

<sup>4</sup>En este informe, el ministerio realizó la estimación del trabajo infantil en el 2011, encontrando que para ese año es de 15 %. La magnitud del trabajo infantil proviene de preguntas realizadas por el DANE, las cuales comprenden dos definiciones: la de ocupación, la cual le hace al individuo las siguientes cuatro preguntas: i) En qué actividad ocupó la mayor parte del tiempo la semana pasada ii) Además de lo anterior, ¿realizó la semana pasada alguna actividad paga por una hora o más? ii) Aunque no trabajó la semana pasada por una hora o más en forma remunerada, ¿tenía durante esa semana algún trabajo o negocio por el que recibe ingresos? y, iv) ¿trabajó o colaboró la semana pasada en un negocio familiar por una hora o más sin que le pagaran? La segunda definición es la de oficios en el propio hogar, la cual establece que un niño, niña o adolescente trabaja si realiza oficios en el propio hogar por más de 15 horas a la semana. Sin embargo, cuando se tiene en cuenta la definición de ocupación esta baja a un 10 %.

los hogares representando un costo de oportunidad. Finalmente, se presentarán diferentes investigaciones que abordan la manera en qué el PAE impacta variables importantes en los niños, niñas y adolescentes.

La literatura es numerosa al abordar cómo las condiciones de la niñez son primordiales a la hora de pronosticar partes importantes del desarrollo como adultos (Feldman, 2011, Durrant, 2011). Sin embargo, es preciso distinguir entre varios tipos de habilidades, estas van desde el desarrollo físico hasta el socio emocional, pero lo interesante de esta parte de la literatura es que encuentra cómo cada una de estas herramientas se une para impactar el futuro de los niños. Por ejemplo, sin una correcta nutrición, los niños y niñas no podrán alcanzar su máximo potencial físico (Grantham-McGregor et al., 2007), pierden la capacidad de concentración y son menos productivos, se incrementan las probabilidades de enfermedades por infección y hasta produce la muerte en casos extremos (Grantham-McGregor et al., 1997; González et al., 2009; Behrman et al., 2003). Así mismo, la alimentación es fundamental a la hora de desarrollar habilidades motoras de los niños; cuando entienden lo que producen y alcanzan a través de sus movimientos, también aprenden a interactuar logrando un desarrollo total e integral (Piaget, 1962 y 1983; Frankenburg et al., 1992; Kail, 2004). Sin embargo, hay ciertas situaciones en las que se ve comprometido este desarrollo integral; la violencia, el entorno y el trabajo infantil sólo para mencionar algunas (Camacho, 2008; Sanchez y Rodriguez, 2012; Heady, 2003). Por lo tanto, un programa de tipo nutricional como el Programa de Alimentación Escolar, no sólo implicaría mejorar los aspectos físicos de los niños, sino que también se espera encontrar que estos contribuyan al desarrollo integral y psicosocial en los niños y las niñas.

Ahora, que los niños y las niñas desarrollen integralmente sus habilidades también garantiza ventajas en su vida adulta. Se ha encontrado que niños bien nutridos y desarrollados, reportaban mejor desarrollo cognitivo entre los 8 y 12 años, lo cual está relacionado con mejores salarios a los 18 y 22 años (Walker et al., 2005; Gertler et al., 2012). En efecto, resultados en pruebas cognitivas están correlacionados con los salarios futuros de personas en edad adulta (Neal y Johnson, 1996, Murnane et al, 1995). Además, otros autores presentan evidencia de que lo acumulado de capital humano hasta los 14 años es más importante en determinar el desempeño laboral posterior de un individuo, que lo que se hace después de esta edad (Keane y Wolpin, 1997; Cameron y Heckman, 1998). Por lo tanto, que el trabajo infantil reduzca los logros educativos de los niños y niñas no sólo determinaría el desarrollo de su educación sino también la calidad de su vida futura (Heady, 2003; Akabayashi y Psacharopoulos, 1999; He, 2016).

En cuanto al capital humano, en línea con la hipótesis de esta investigación, la segunda parte de la literatura se enmarca dentro de la teoría del costo de oportunidad de la educación. En ese sentido, Becker (1964) plantea que el gasto en educación incluye la inversión monetaria y el costo de estar en la población económicamente inactiva, tales como las rentas que no se reciben por estudiar. Además, el autor propone un modelo teórico donde la familia es una unidad en la que se toman decisiones eficientes, de tal forma que cada miembro se especializa en labores domésticas o en el mercado laboral, de acuerdo a las ventajas comparativas. Por lo tanto, la decisión de la familia de enviar a los hijos a estudiar, en vez de trabajar, es una decisión económica racional que depende del costo oportunidad. Este argumento se convierte en teoría central de esta investigación, debido a que para los niños y sus familias estar en el colegio representa un costo compuesto por el valor de la matrícula, útiles escolares, comida necesaria para tener energía y por tiempo invertido. Ahora, ya que la comida se incluye dentro de los gastos escolares, el PAE, al ser una transferencia en especie, ayuda a disminuir el presupuesto que las familias invierten en alimentación, es decir, es capaz de reducir el costo de oportunidad de estar en el colegio.

A partir del modelo de Becker (1964) que plantea que existe un costo de oportunidad también determinado por estar por fuera del mercado laboral, se han generado varios modelos teóricos específicos entre el trabajo infantil y la educación. Baland y Robinson (2000) construyen un modelo con el fin de encontrar las implicaciones del trabajo infantil sobre el bienestar de los niños. El resultado es que existe un trade-off entre la participación laboral y la acumulación de capital humano, explicado por la inversión eficiente de los padres en la educación de sus hijos. Ahora bien, en términos empíricos, Ravallion et al (2000) presentan evidencia de que el trabajo infantil se genera a expensas de la educación, encontrando que la participación laboral temprana se produce cuando la familia del niño ya no considera rentable enviarlo al colegio. Por su parte, Akabayashi y Psacharopoulos (1999) muestran que existe sustitución entre horas de trabajo y de estudio, de tal manera, que los niños que trabajan tienen peores resultados en matemáticas y lectura, en comparación con los que no trabajan.

Dichas investigaciones soportan la idea de costos de oportunidad atados a la educación y la manera en que el trabajo infantil aprovecha esta situación para manifestarse; de todas formas las familias ven la educación de sus hijos como una decisión de inversión, y cuando ésta ya no representa una rentabilidad, los hogares dejan de enviar a sus hijos al colegio y en vez de eso los mandan al mercado laboral, es decir, reemplazan asistencia escolar por trabajo infantil. A esto se suma, que existe una relación negativa entre trabajo infantil y acumulación de capital humano.



Teniendo en cuenta la importancia del desarrollo de habilidades en la niñez y la adolescencia, y el costo de oportunidad de la educación, la tercera parte de la literatura se enfoca en entender cómo se generan y cómo se desarrollan desde la primera infancia hasta los 14 años estas cualidades mediante programas como el PAE. En efecto, el ánimo de encontrar la mejor fórmula para desarrollar en los niños estas capacidades ha motivado y justificado la generación de intervenciones desde la primera infancia y hasta la edad escolar, estrategias en donde se pueden enmarcar los programas de alimentación escolar.

Al respecto, aunque existe literatura a nivel internacional que no encuentra efectos significativos de este tipo de programas para países como Filipinas, India y Chile (Respectivamente Jacoby, 2002; Afridi, 2010; McEwan, 2013), hay investigaciones en este mismo espectro que sustentan el impacto del PAE a nivel nutricional y escolar. En efecto, algunas investigaciones encuentran impacto a nivel nutricional, debido a que el programa contribuye a reducir los niveles de anemia (Alderman y Bundy, 2011), mejorar los índices de micronutrientes (Van Stuijvenberg et al, 1999), controlar la obesidad e incluso promover una buena higiene (Bundy et al, 2006). Además, es de notar que estos programas generan una especie de reasignación eficiente de comida a otros miembros, beneficiando a los hermanos menores de los niños receptores de PAE, mostrando efectos superiores sobre estos a nivel nutricional (Islam y Hoddinot, 2009; Kazianga et al, 2009). Que este programa beneficie a los hermanos menores podría explicar el hecho de que algunas investigaciones no encuentren efectos significativos provenientes del PAE, ya que puede que no se estén enfocando en las poblaciones que más aprovechan estos programas como son los niños menores, que como ya se comentó, son capaces de aprovechar mejor las intervenciones. Este documento va más allá de variables que midan los efectos nutricionales del programa en los niños y las niñas, más bien el propósito de este es observar la manera en qué impacta la seguridad de la familia y por consiguiente sus decisiones, así mismo, ya se discutió cómo el trabajo infantil es capaz de obstaculizar la integralidad y por tanto la vida futura de los niños, por lo que también afectaría indirectamente estas variables.

Así como esta investigación indaga en los efectos del PAE en el trabajo infantil, también es reconocido el impacto del programa en variables más de tipo social como la educación; de hecho son numerosos los estudios que demuestran sus efectos en asistencia, especialmente en áreas con indicadores de participación escolar bajos (Jukes et al, 2007; Kristjansson et al, 2007; Adelman et al, 2008). Ahora, en términos de rendimiento escolar los estudios no parecen ser concluyentes; por ejemplo, mientras el programa puede mejorar la atención de los niños, su impacto en el aprendizaje

también depende de otros factores como la organización de la clase (Grantham-McGregor et al, 1998), la calidad de la educación (Whaley, 2003) y los efectos pares generados por estudiantes con no tan buenas habilidades (Ahmed et al, 2006).

También existe evidencia que soporta el impacto de estos programas en América Latina y el Caribe. En Perú, a pesar de que el programa no tuvo efecto en la altura para la edad ni en el índice de masa corporal, se reportan aumentos en la memoria de corto plazo y los puntajes en matemáticas y comprensión de lectura, así como una tasa de asistencia superior comparando con los niños por fuera del programa (Cueto et al, 2000). En el caso de Jamaica, el programa, además de aumentar la asistencia escolar, mejoró la altura para la edad y se encuentra un efecto positivo para las pruebas de matemáticas de los niños entre segundo y tercer grado (Powell et al, 1998). Para Colombia, en el 2013 se publicó una Evaluación de Operaciones y Resultados para determinar el grado de Efectividad del Programa de Alimentación Escolar (CINDE, 2013). En dicha evaluación, se encuentra que la aceptación del PAE por parte de los beneficiarios, y de cada uno de los distintos complementos alimentarios, es muy alta (cercana al 100%). Así mismo, se encuentra que existe una buena focalización en tanto, al servir la alimentación escolar en establecimientos oficiales, llega a los niños y adolescentes más pobres del país. Dentro de este estudio, también se realiza una evaluación de impacto y se encuentra que el programa contribuye a disminuir la deserción escolar, especialmente en las zonas rurales.

Finalmente, a la hora de analizar el trabajo infantil para Colombia y su relación con programas sociales, Attanasio y Gómez (2004) documentan el efecto del programa Familias en Acción en varios resultados de la familia como el trabajo de los niños, de la madre y variables del estado de salud. El autor encuentra efectos heterogéneos entre el área rural y urbana, siendo Familias en Acción más eficiente en el sector rural en cuanto a la reducción del trabajo infantil, el aumento de la ocupación de las madres y en variables de salud. Por su parte, Quiroga (2006) se enfoca en trabajo infantil y evalúa empíricamente la decisión de los beneficiarios del programa Familias en Acción de enviar a sus hijos a estudiar o a trabajar, utilizando el método de emparejamiento. La autora encuentra que el programa reduce el trabajo infantil y aumenta la tasa de asistencia de los niños trabajadores. Dichos resultados son compatibles con el propósito de esta investigación ya que la hipótesis es que así como familias en acción reduce el trabajo infantil, el Programa de Alimentación Escolar disminuye este, ya que ambos programas son capaces de cubrir parte del costo de oportunidad de enviar al niño al colegio.

### 3. Marco Teórico

Ahora bien, teniendo en cuenta la literatura expuesta acerca de lo costosa que puede llegar a ser la educación y los aportes modestos pero existentes del Programa de Alimentación Escolar, es preciso también exponer la manera en cómo y por qué estos programas pueden reducir el trabajo infantil, exposición que servirá como marco teórico para esta investigación.

En ese sentido, un aspecto a considerar es la manera en que el ambiente y los factores socio económicos pueden generar trabajo infantil. Un aspecto fundamental es la pobreza, ya que como bien lo dijo Becker (1964), al ser las familias unidades en las que se toman decisiones, el hogar por falta de recursos económicos puede considerar más rentable enviar al niño al mercado laboral que al colegio. De ahí que sea claro que la pobreza es un factor generador del trabajo infantil. Esta relación ha sido encontrada por diversos autores y para un amplio conjunto de países como India, Pakistan, Ghana y Egipto (Respectivamente, Rosenzweig, 1981 y Labenne, 1997; Ray, 2000; Canagarajah y Coulombe, 1998; Wahba, 2001). En Latinoamérica se encuentran trabajos como el de Kruger en la parte rural de Brasil (2004), donde encuentra que la probabilidad de que un niño trabaje aumenta en hogares de ingresos bajos y medios, mientras que no pasa nada en las familias de ingresos altos. Para Colombia, Rodríguez y Sánchez (2012) relacionan el trabajo infantil con trampas de pobreza, que se ven agudizadas cuando un niño abandona el colegio por entrar en el mercado laboral. Que consistentemente se encuentre literatura que demuestra que la pobreza es capaz de producir trabajo infantil, es señal de que programas sociales dirigidos a población vulnerable, que consecuentemente reducen la pobreza, son capaces también de neutralizar los efectos colaterales de esta. Sin embargo, es preciso analizar primero más formas en que a través de la pobreza cobra vida el trabajo infantil.

En primer lugar, se ha demostrado que choques negativos en el ingreso aumentan la deserción escolar y el trabajo infantil. Guarcello et al (2010) encuentran un progreso considerable en términos de la reducción del trabajo infantil explicado por mejoras en la calidad de vida en Guatemala. Dicha hipótesis se mantiene para India, Indonesia y Tanzania (Respectivamente Jacoby y Skoufias, 1997; Thomas et. al., 2004; Beegle y Weerdt, 2006). Lo anterior, sugeriría que no sólo una variable de pobreza de tipo estado podría generar trabajo infantil, sino que también cambios en la riqueza de los hogares también dan lugar a esto, precisamente los programas o intervenciones pueden ser considerados como generadores de choques, por lo tanto si el PAE es capaz de generar cambios en

la situación de riqueza o bienestar entonces podría estar cambiando el trabajo infantil.

Adicionalmente, buena parte de la literatura se especializa en analizar cómo el trabajo infantil se concentra en zonas más vulnerables como las rurales y en condiciones económicas precarias. Dammert (2005) encuentra que el trabajo infantil y el ingreso de los hogares peruanos tienen una relación negativa. De este modo, los aumentos en el ingreso, incrementan la probabilidad de que niños y niñas estudien. En particular, el autor agrega, que en las zonas rurales la reducción en el trabajo infantil es mayor, en comparación con la zona urbana, cuando se incrementa el ingreso familiar. Así mismo, es interesante para esta investigación lo registrado por Ersado (2005), quien realiza comparaciones de trabajo infantil entre las áreas urbanas y rurales en Nepal, Perú y Zimbabwe, concluyendo que la pobreza reduce la probabilidad de asistencia al colegio e incrementa la prevalencia e intensidad del trabajo infantil, medido en horas, encontrando que los rurales están en situación más precaria. Los resultados de Ersado (2005), son consistentes por los hallados con esta investigación que se propone encontrar que el PAE reduce el trabajo infantil y mejora el bienestar de los hogares, lo que se evidencia en una reducción de la tasa de trabajo infantil. En línea con Ersado (2005), para observar si los mecanismos son los correctos, este documento también muestra aumentos en asistencia escolar y actividad física de los niños, lo que supone que los niños están reemplazando trabajo infantil por otras actividades imprescindibles para su desarrollo integral.

Luego de mostrar la relación entre pobreza y trabajo infantil, se debe considerar el canal en que el Programa de Alimentación Escolar, al ser una transferencia en especie a los hogares, contribuye a aumentar la riqueza de los hogares, en la medida en que ahora los padres no tienen que dedicar dinero para los almuerzos o desayunos de sus hijos. Por lo tanto, si el PAE reduce la pobreza de los hogares expuestos, a su vez, estos tendrán menor probabilidad de enviar a sus hijos al mercado laboral.

Al respecto, Alderman y Bundy (2011) analizan evidencia reciente y encuentran que los Programas de Alimentación Escolar son vistos como una transferencia de ingresos dirigida a hogares con pocos recursos, que además tienen el beneficio adicional de generar capital humano mejorando la situación de salud de los receptores. Dichos objetivos hacen que estos programas se conviertan en herramientas para reducir la pobreza, mejorando el bienestar de los hogares y evitando programas asistenciales y transferencias en un futuro. También Das et al (2005) explican que este tipo de programas implican eficiencia y equidad, generan externalidades positivas y alivian fallas de mercado, al tiempo que combaten la desigualdad, esto cuando redistribuyen los recursos hacia los

más pobres.

Por lo tanto, esta investigación hará uso de esas experiencias y se centrará en la idea de que los programas de Alimentación Escolar mejoran el bienestar de los hogares alejando a los niños del trabajo infantil y acercándolos a la escuela. Esto también permitirá demostrar que en el caso colombiano son los hogares más pobres los que presentan mayor proporción de participación laboral infantil. Esta idea va en contra de lo elaborado por Bhalotra y Heady (2003) quienes encuentran para Pakistán y Ghana que las niñas en hogares con mayor acumulación de tierra tienden a trabajar más, lo que podría estar sugiriendo que la reducción del trabajo infantil no se debe al programa sino a cambios en la riqueza de los hogares<sup>5</sup>. Para evaluar lo anterior, se hará uso de un índice de seguridad alimentaria que se aproxime al componente de pobreza de las familias. Así, la definición usada por la ENSIN (2010) considera que esta es la disponibilidad de alimentos de manera suficiente y estable, así como un acceso oportuno y permanente de los mismos, en términos de cantidad, calidad e inocuidad, componentes que podrían estar también reflejando la capacidad adquisitiva de los hogares. De hecho, esta variable originalmente es una escala compuesta por varios ítems que cubren la preocupación de los hogares por falta de alimentos, recursos para comprarlos, reducción de la disponibilidad de comida y la experiencia del hambre. Es una escala que mide la condición de hogar como grupo y que va de 1 a 4 donde la variable fue recodificada para que 1 fuera inseguridad severa y 4 el más alto grado de seguridad. La escala refleja el proceso teórico de la inseguridad alimentaria; comienza como un estímulo negativo que genera preocupación, pero si no es tratada correctamente puede llevar a experimentar hambre (Encuesta Nacional de Situación Nutricional, 2010)<sup>6</sup>.

Para estar seguros de si esta variable puede ser una proxy de riqueza de los hogares, es de notar la variedad de investigaciones para Colombia, Latinoamérica y Asia que han demostrado que los

---

<sup>5</sup>Para resolver esto, también es de notar que en esta investigación se hace uso de controles de estrato y Sisbén como proxys de la pobreza, de hecho la encuesta Sisbén recoge toda la información referente a la riqueza. En particular, dicha encuesta indaga sobre el origen de los ingresos, incluyendo la opción rentista cuando la persona tiene posesiones o inmuebles que rinden alguna renta

<sup>6</sup>Las 4 categorías que integran la escala y que fueron recodificadas para propósitos de esta investigación son: 1) Inseguridad severa: todos los miembros del hogar han reducido la ingesta de alimentos y hasta los niños experimentan hambre; 2) Inseguridad moderada: la ingesta en los adultos fue reducida y estos experimentan hambre; 3) Inseguridad leve: se evidencia preocupación en el hogar por el abastecimiento suficiente de alimentos y se reporta disminución en la calidad de estos; 4) seguridad alimentaria: el hogar muestra nada de inseguridad alimentaria. También cabe resaltar que para los resultados expuestos en cuanto a los impactos del PAE en seguridad alimentaria se hace uso de una variable dummy construida sobre la lógica de que 1 es seguro y 0 es inseguro, dada la definición de las variables anteriores.

hogares más expuestos a la inseguridad alimentaria son los encabezados por campesinos pobres con producción de alimentos marginal o inadecuada, que están compuestos por un alto número de personas, que viven en áreas riesgosas y por consiguiente con ingresos muy bajos (Álvarez et al 2007, Melgar-Quinonez et al, 2006 y Hackett et al, 2010). Estas investigaciones soportan la idea de que la seguridad alimentaria integra el componente de riqueza de los hogares, variable que además refleja la capacidad adquisitiva para comprar y disponer de alimentos. Además, es sumamente valioso que la variable sea capaz de recoger las percepciones de los hogares con respecto a su alimentación, por lo que sería una excelente variable mediadora a la hora de capturar cómo las decisiones familiares son afectadas por un programa, precisamente, de tipo nutricional. Por lo tanto, es preciso analizar cómo la literatura aborda la relación de la seguridad alimentaria y las decisiones de los hogares.

Así como la inseguridad alimentaria tiene implicaciones físicas también genera cambios sociales que afectan el núcleo familiar; hambre, depresión, enfermedad y bajo rendimiento escolar son algunos factores asociados a esta situación. Al mismo tiempo, estas perturbaciones son capaces de cambiar los patrones de consumo de la familia, afectando la calidad y cantidad de la ingesta diaria e inclusive induciendo a la utilización de formas no convencionales de adquirir alimentos. El hurto, la prostitución, pedir limosna y el mismo trabajo infantil pueden convertirse en los nuevos mecanismos de los hogares para conseguir comida (Hameli et Al 1999; Webb et al 2006). Dichas razones hacen que esta investigación utilice la seguridad (inseguridad) alimentaria como una variable que explique la diferencia que ocurre en las familias, cuyas decisiones y percepciones varían cuando el programa de alimentación escolar les brinda cierto grado de seguridad.

Ahora según el marco conceptual presentado por Alderman y Bundy (2012), quienes consideran que los Programas de Alimentación Escolar necesitan ser repensados a la hora de estimar sus efectos, esta investigación considera que no es correcto creer que un programa de bajo costo, como el PAE <sup>7</sup>, afecte índices de riqueza convencionales, que simplemente determinan patrones de riqueza absoluta y que no van más allá de contar los bienes y servicios de las personas. Por lo tanto, como el indicador de seguridad alimentaria hace un esfuerzo por capturar las capacidades y limitaciones tanto físicas como mentales del hogar que están determinadas por su alimentación,

---

<sup>7</sup>En el 2010, el costo promedio de un almuerzo escolar era cercano a las 1200 pesos, mientras que un desayuno costaba cerca de 900 pesos. Para el 2015, el Ministerio de Educación acordó como costo promedio por ración 943 pesos para desayunos y 1.282 pesos para almuerzos, pero los precios mínimos calculados y revisados determinaron que los primeros debían costar 1.370 pesos y los segundos de 2.273 pesos (Anexo Técnico Programa Alimentación Escolar, 2010)

es claro que el PAE puede tener repercusiones inmediatas en la seguridad alimentaria o lo que es lo mismo, en la disponibilidad de riqueza de los hogares para alimentación o capacidad de consumo. En ese orden de ideas, sí el PAE hace que las familias se sientan más seguras, y como se analizó que el trabajo infantil se presenta en esas situaciones de inseguridad, consecuentemente, estos hogares más seguros serán capaces de tener a un niño fuera del mercado laboral.

## **4. Datos y Estadísticas Descriptivas**

Para efectos de esta investigación se utiliza la Encuesta Nacional de Situación Nutricional (ENSIN), para el año 2010. La ENSIN tiene cobertura nacional con representatividad urbana y rural, para seis regiones y 32 departamentos. La base de datos abarca la población colombiana entre 0 y 64 años de edad, y la información disponible se presenta desagregada por grupos de edad, sexo y nivel socio económico. La encuesta cuenta con 8 módulos o componentes,<sup>8</sup> pero sólo se emplean los módulos que incluyen datos Antropométricos, Actividad física y el de seguridad alimentaria el cual cuenta con información para niños en edad escolar.

Es importante entrar a discutir la variable dependiente usada en esta investigación, en este caso para capturar el trabajo infantil se toma la pregunta ¿Trabajó la semana pasada? Dicha pregunta se encuentra enmarcada dentro de la metodología expuesta por el Ministerio de Trabajo y está en línea con la definición de la OIT. La definición de trabajo infantil no depende de si el niño recibe remuneración, el sólo hecho de que el niño utilice su tiempo en trabajar quiere decir que está interfiriendo con su evolución física y mental, es decir, está usando su tiempo para trabajar y no para hacer tareas, estar en el colegio, o desarrollarse física y mentalmente. Vale la pena notar que en este caso no se está tomando en cuenta los oficios realizados en el hogar, ya que como se mencionó anteriormente, estas actividades son motivo de protección, más no de erradicación, por lo tanto desde la teoría no sería deseable ningún efecto. Así mismo, a través de la ENSIN se calcula que la tasa de trabajo infantil en Colombia es alrededor de 9 %. Dicha cifra concuerda con la cifra oficial que fue hallada por el Ministerio de trabajo en el 2011, cuya metodología ya fue expuesta anteriormente en este documento (Ministerio de Trabajo, 2011).

---

<sup>8</sup>Los 8 componentes son: Antropométricos y Bioquímicos, Lactancia materna y alimentación complementaria, Seguridad alimentaria en el hogar, Prácticas de alimentación de interés en nutrición, Actividad física, Tiempo dedicado a ver televisión o jugar con video juegos y Auto percepción de peso corporal.

Como se observa en el cuadro 1 se utiliza una muestra de los niños entre los 5 y 17 años de edad (47.583 observaciones) quienes son la población objetivo del PAE. De acuerdo con la base de datos el 52 % de los niños tiene el programa de alimentación escolar (PAE), cifra de cobertura que a su vez es manejada de manera oficial (aproximadamente 4 millones de niños) (ICBF, 2012).

Ahora bien, es importante analizar las características de la muestra a nivel de grupo de control y tratamiento de tal manera que se puedan observar los desbalances antes de aplicar la metodología. El cuadro 2 presenta las variables socio demográficas dividido por grupos conformados por niños que reciben PAE (tratados) y los que no (controles). Se observa que existen diferencias importantes entre el grupo de tratamiento y control. En particular, los niños con PAE provienen de hogares más vulnerables, menor Sisbén<sup>9</sup>, son más jóvenes, cursan un menor grado, tienen mayores probabilidades de tener a sus padres en el hogar, tienen hogares más numerosos, menor probabilidad de tener una vivienda propia, menor probabilidad de que el cabeza de hogar perciba ingresos, sus madres son menos educadas y con mayor probabilidad de vivir en el área rural, entre otras.

Por otro lado, la figura 1 muestra cómo los niños sin PAE son los que más trabajan comparados con los niños con PAE. Así mismo las figuras 2, 3 y 4 muestran cómo este mismo fenómeno se presenta en el sector rural, en los niños en secundaria y en los hombres, en donde se aprecia que los niños sin PAE son los que efectivamente trabajan más.

Es importante mencionar que se trabajó sin excluir a los niños que no asisten al colegio, ya que dentro de los niños que trabajan el 64 % también asiste a un plantel educativo. Además, si bien para recibir PAE un niño debe estar en el colegio, asistir a la escuela no garantiza que recibirá PAE, ya que en el 2010 sólo el 52 % de los colegios tiene PAE (como se menciono anteriormente). De igual manera, al incluir asistencia como una variable que explica PAE muestra problemas de colinealidad, por lo que no se incluirá en las estimaciones. Sin embargo, la probabilidad de tener PAE no sólo se determina por asistir al colegio, porque también depende de otras variables socio económicas.

A su vez, esta investigación también hace uso de una variable como proxy de la riqueza de los hogares, en este caso seguridad alimentaria, con la que además de comprobar el mecanismo planteado se confirma que en Colombia no se aplica lo encontrado por Bhalotra y Heady (2003), quienes encuentran que el trabajo infantil se puede presentar también para hogares con mayor

---

<sup>9</sup>Esto es de esperar porque una de las formas de priorizar a la población que recibe el PAE es utilizando la clasificación más baja del Sisbén.



riqueza. Por eso, en el cuadro 3 se presentan las estadísticas descriptivas relacionadas con este índice y el promedio de esta a través de otras variables socio económicas como son el área urbana y rural, la participación laboral infantil y si es beneficiario de PAE. Lo primero que se aprecia es que efectivamente, las personas que habitan en áreas rurales tienen mayor inseguridad alimentaria, esto se mantiene para los niños con PAE, que en promedio son más vulnerables que los niños que no son beneficiarios de este. Ahora, lo mismo ocurre para los niños y niñas que trabajan en todo el país, ya que los que están en el mercado laboral están en promedio en un nivel de inseguridad más severo que los que no trabajan. Al mismo tiempo, observando sólo a la población del sector rural se mantiene que los niños y niñas que trabajan en estas zonas son más inseguros que los que no lo hacen. Esto es importante en esta investigación, ya que se abordará el impacto del PAE discriminado por áreas y mostrando además que para Colombia no permanece lo expuesto por Bhalotra y Heady (2003). Para efectos de esta documento, también es preciso mostrar lo que ocurre con los niños que tienen PAE y que trabajan, que también experimentan mayor inseguridad dentro de los niños que reciben el programa.

## 5. Estrategia Empírica

El objetivo de esta investigación es encontrar la relación entre el PAE y el trabajo infantil. Con este fin, el primer modelo que se plantea es estimar la siguiente ecuación con mínimos cuadrados ordinarios (MCO):

$$Y_{ij} = \alpha + \beta_1 PAE_i + \beta_2 X_i + \eta_j + \varepsilon_{ij} \quad (1)$$

Donde  $Y_{ij}$  es una dummy igual a 1 si el niño  $i$  trabaja en el municipio  $j$ .  $PAE_i$  es igual a 1 si el niño recibe el programa de alimentación escolar,  $X_i$  son características socio demográficas, es decir, todas las presentadas en el cuadro 1 y  $\eta_j$  son efectos fijos del municipio donde vive el niño que controlan por variables observadas y no observadas que sean iguales para todos los niños dentro de un municipio,<sup>10</sup>. Finalmente,  $\varepsilon_{ij}$  es el término de error.

---

<sup>10</sup>Una variable a controlar gracias a estos efectos, son los operadores encargados de distribuir y preparar los alimentos para los niños

Sin embargo, el modelo de mínimos cuadrados ordinarios (MCO) puede calcular estimadores sesgados<sup>11</sup>. Primero, porque no controla el problema de endogeneidad que surge de que los niños con PAE son más vulnerables y por tanto más expuestos a trabajo infantil. En efecto, las estadísticas descriptivas reflejan esta endogeneidad, ya que reportan diferencias significativas entre el grupo de tratamiento y el grupo de control poniendo en evidencian que los dos grupos, son distintos en dimensiones relevantes para el análisis de los efectos del programa.

Así mismo, es probable que los participantes hayan sido auto-seleccionados en el programa con base en características observadas o no observadas. Por ejemplo, una variable no observada es la habilidad de las madres para hacer que sus hijos sean seleccionados dentro del programa de alimentación escolar, es decir, se pueden hallar niños que sean intervenidos gracias a que sus madres son sagaces y los matriculen en un colegio donde se brinde el PAE. Lo anterior, generaría que los niños participantes sean los que tienen madres que tal vez trabajen más y que no necesiten enviarlos al mercado laboral, o incluso mamás con un horizonte de planificación amplio capaces de alejar a sus hijos del trabajo infantil porque entienden las nefastas consecuencias que este les acarrea. En dicho caso, MCO arrojaría coeficientes de mayor magnitud a los de un método que controle por aquellas características de niños y hogares que determinan tanto la probabilidad de participar en el programa como la probabilidad de trabajar de los niños.

Para solucionar el problema de MCO, en primer lugar se desarrolla un modelo de emparejamiento o Propensity Score Matching (PSM). La idea de usar esta metodología es que cuando se controla por características observadas como educación de la madre, presencia del padre en el hogar, nivel de Sisbén, municipio, género y edad del niño, entre otros, la decisión de participar en el programa se determina por este conjunto de variables y al mismo tiempo se logra controlar por las no observables. Luego de controlar por estas características, esta metodología encuentra un clon dentro del grupo de control para cada niño del grupo de tratamiento (con PAE), donde el clon es el individuo lo más similar posible en variables observadas al niño que recibe PAE.

El PSM se calcula en dos pasos. Primero, se calcula Propensity Score (PS) o la probabilidad de ser tratado con la siguiente ecuación estimada mediante un modelo logit.

$$PAE_i = \alpha + \beta_1 X_i + \eta_j + \varepsilon_{ij} \quad (2)$$

---

<sup>11</sup>Los resultados en la siguiente sección muestran que mínimos cuadrados ordinarios sobre-estima el efecto del PAE en el trabajo infantil comparado con las técnicas de emparejamiento

Segundo, se emparejan un individuo de tratamiento (con PAE) con uno de control, usando el método de PSM, planteado por Rousenbaum et al (1983) y que describe el estimador de emparejamiento como:

$$E[Y_{1i} - Y_{0i} | D_i = 1] = \sum_x \theta_x P[(X_i = x) | D_i = 1] \quad (3)$$

Donde  $Y_{1i}$  es el resultado potencial de los tratados y  $Y_{0i}$  es el resultado potencial de los controles,  $D_i$  es una dummy igual a 1 cuando el niño recibe el PAE,  $X_i$  es un conjunto de características socio económicas del individuo  $i$  y  $P[(X_i = x) | D_i = 1]$  es la probabilidad de recibir el tratamiento dependiendo de características individuales ( $X_i$ ) (Propensity Score). Este último término sirve para ponderar finalmente a los individuos que son emparejados mediante el parámetro  $\theta$ . De esta manera, se puede interpretar dicha estimación como un MCO ponderado por la probabilidad de pertenecer al programa, y se halla el efecto promedio del tratamiento sobre los tratados ( $\sigma_{ATT}$ ), es decir, el efecto del PAE sobre los niños entre 5 y 17 años que trabajan.

Ahora existen varios algoritmos de emparejamiento como son vecino más cercano, kernell, distancia de Mahalonobis y por estratificación (Bernal y Peña, 2011). Siguiendo a Quiroga (2006), Bernal (2013) y otros autores que han impletado propensity score matching, se utilizará el emparejamiento por vecino más cercano<sup>12</sup>.

El emparejamiento por vecino más cercano empareja a cada individuo tratado con el de control más similar, es decir, busca entre los controles un agente con la probabilidad de participación más cercana o lo que es lo mismo el vecino más cercano con base a  $P(X)$ . Ahora, el vecino de control más cercano,  $C(i)$ , del individuo de tratamiento  $i$  está dado por:

$$C(i) = \{j \in D = 0 | \argmin_j ||P_i(X) - P_j(X)||\} \quad (4)$$

Esta ecuación quiere decir que se elige al individuo en el grupo de control tal que la distancia entre su probabilidad predicha de participación y la del individuo del tratamiento sea mínima. Ahora, el impacto promedio del programa sobre los tratados (ATT) de vecino más cercano está dado por:

---

<sup>12</sup>Sin embargo, en los resultados serán usados los efectos encontrados usando Kernell

$$\tau_{ATT}^{PSM-VC} = \sum_{i=1}^I \text{Promedio}((Y_i|D_i = 1) - (Y_{c(i)}|D_i = 0)) \quad (5)$$

Donde  $i$  es un individuo que pertenece al grupo de tratamiento,  $I$  es el número total de personas que reciben PAE,  $(Y_i|D_i = 1)$  es la variable de resultado para el individuo del grupo de tratamiento,  $C(i)$  es el conjunto de vecinos más cercanos en el grupo de control para el agente y en ese orden de ideas,  $(Y_{c(i)}|D_i = 0)$  es el trabajo infantil para cada individuo más cercano a  $i$ .

Si bien esta metodología parece lógica a la hora de aproximarse al verdadero coeficiente y es más útil que una simple comparación de medias (porque elimina el sesgo de los estimadores), es importante mencionar que el método tradicional de Propensity Score Matching presenta retos metodológicos. Smith y Todd (2005), por ejemplo, encuentran que los resultados del emparejamiento son sensibles a la inclusión de variables y a cambios en la muestra. Así mismo, Arceneaux (2005) presenta evidencia de que el emparejamiento produce resultados sesgados, incluso utilizando un gran número de características observadas.

A pesar de las desventajas del emparejamiento, recientemente se han desarrollado modelos que controlan los problemas mencionados y producen estimadores insesgados. McKenzie et al (2010) muestran que el emparejamiento con corrección de sesgo logra conseguir resultados casi idénticos que los de un experimento natural<sup>13</sup>. Dicha investigación concluye que resultados más creíbles pueden ser alcanzados mediante métodos no experimentales si el tratamiento y el grupo de control tienen un amplio soporte común y las personas entrevistadas se les administra una misma encuesta, supuestos que se mantienen en esta investigación (Dehejia y Wahba, 2002; Smith y Todd, 2005).

Abadie e Imbens (2004, 2005 y 2011) se han encargado de estudiar y mejorar el método original del emparejamiento creado por Rosenbaum y Rubin (1983). Los autores presentan una aproximación de regresiones multivariadas con corrección de sesgo, que a lo largo de la literatura es destacada como creíble en términos de identificación, ya que es capaz de recuperar los efectos causales de los programas. Básicamente, esta metodología considera el mismo problema de sesgo que plantea el emparejamiento simple y del que adolece MCO: una simple comparación de promedios puede ser contaminada por el efecto de otras variables que están correlacionadas con el

---

<sup>13</sup>Mckenzie et al (2010) muestra que el orden de las metodologías que más aproximan sus resultados a un experimento natural son las variables instrumentales, seguido por el matching con corrección de sesgo y la diferencia en diferencias.

tratamiento  $Di$  <sup>14</sup> como también con los resultados potenciales  $Yi(1)$  y  $Yi(0)$ . Dado que la existencia de estas variables puede crear una correlación entre  $Di$  y  $Yi$  <sup>15</sup>, los autores exponen una serie de supuestos para la identificación y estimación del efecto promedio del tratamiento. En resumen, los supuestos van desde condiciones en el soporte de los covariados hasta los supuestos tradicionales de independencia condicional y soporte común <sup>16</sup>.

En particular, Abadie e Imbens (2011) hacen una versión del estimador típico de emparejamiento bajo la idea de estimar el resultado potencial que no es observable. En esencia, este método de emparejamiento con corrección de sesgo imputa el resultado potencial faltante para cada sujeto, lo anterior usando un promedio de los resultados de sujetos similares que están expuestos al tratamiento contrario. En este caso, el efecto promedio de los tratados (ATT) es computado tomando el promedio de la diferencia entre el resultado observado y el resultado potencial imputado <sup>17</sup>.

Ahora, bajo la idea de encontrar el resultado no observado mediante imputación, los autores consideran un estimador del emparejamiento con corrección de sesgo, en que la diferencia con un modelo estimado por MCO es que la regresión de imputación va a ser ajustada por la diferencia de la función de regresión estimada en valores de los covariados para  $i$  y la función de regresión estimada en los covariados de las parejas  $j$ , a saber:

$$\tilde{Y}_i(0) = \begin{cases} Y_i & \text{si } Di = 0 \\ \left(\frac{1}{M}\right) \sum_{j \in M(i)} (Y_j + \hat{\mu}_0(X_i) - \hat{\mu}_0(X_j)) & \text{si } Di = 1 \end{cases} \quad (6)$$

y

$$\tilde{Y}_i(1) = \begin{cases} \left(\frac{1}{M}\right) \sum_{j \in M(i)} (Y_j + \hat{\mu}_1(X_i) - \hat{\mu}_1(X_j)) & \text{si } Di = 0 \\ Y_i & \text{si } Di = 1 \end{cases} \quad (7)$$

<sup>14</sup> $D$  denota si está asignado al tratamiento  $Di=1$  ó al control  $Di=0$ ,  $M$  es el conjunto de parejas utilizadas tal que  $M \leq N_0$  y  $M \leq N_1$ , es decir, el conjunto de parejas  $M$  es menor al grupo de tratados y controles.

<sup>15</sup>Este problema puede ser eliminado a través de una aleatorización

<sup>16</sup>Para una descripción detallada de estos supuestos y de la manera en que se realiza el emparejamiento consulte el apéndice 1

<sup>17</sup>Aunque el estimador original que utilizan Abadie e Imbens en el 2011 es el efecto promedio del tratamiento (ATE), es decir el efecto del programa incluso sobre niños que no lo reciben, ya que uno de los propósitos del documento es hacer una comparación a través de diferentes métodos, se usará la extensión del estimador que funciona con el ATT

Donde  $\hat{\mu}_0(X_i)$  es la función de regresión estimada en los covariados utilizados para el individuo  $i$  en el grupo de control, mientras que  $\hat{\mu}_0(X_j)$  es la función de regresión estimada para las parejas  $j$ . Respectivamente, esto aplica para los individuos tratados en cuanto a  $\hat{\mu}_1(X_i)$  y  $\hat{\mu}_1(X_j)$ .

Finalmente, el estimador resulta de:

$$\tau_M^{bcm} = \frac{1}{N} \sum_{i=1}^N (\tilde{Y}_i(1) - \tilde{Y}_i(0) | D_i = 1)^{18} \quad (8)$$

Ahora bien, ya que se presentó el estimador correspondiente al de emparejamiento con corrección de sesgo, se propone discutir las ventajas metodológicas de este frente a MCO y emparejamiento simple. Primero, la versión construida por Abadie e Imbens (2011) muestra que el estimador del emparejamiento con corrección de sesgo es robusto contra toda mala especificación de la función de regresión, corrigiendo el problema inicial de sesgo denotado por Smith y Todd (2005)<sup>19</sup>. Segundo, los estimadores de emparejamiento con corrección de sesgo son  $N^{1/2}$  consistentes y asintóticamente normales sin importar el número de covariados, al mismo tiempo se comporta mejor en términos de error cuadrático medio y cobertura para intervalos de confianza<sup>20</sup>. Dichas propiedades hacen que este estimador tenga robustez adicional debido a que garantiza consistencia para cualquier valor dado de los parámetros sin necesidad de aproximaciones precisas, y sin tener que lidiar con la sensibilidad del emparejamiento al ajuste por covariados<sup>21</sup>.

Ya que la idea de esta investigación es acercarse a coeficientes insesgados y que no presenten los usuales problemas propios de métodos de emparejamiento convencionales, también se sigue

---

<sup>18</sup>Se considera la versión para el efecto promedio sobre los tratados con el ánimo de hacer comparaciones.

<sup>19</sup>Sea  $\mathcal{J}_M(i) = \mathcal{L}_1(i), \dots, \mathcal{L}_M(i)$  que denota el conjunto de índices para la primera pareja de  $M$  para la unidad  $i$ , para  $M$  tal que  $M \leq N_0$  y  $M \leq N_1$ , es decir, el conjunto de parejas  $M$  es menor al grupo de tratados y controles. Si la función de regresión no es la adecuada, la expectativa de  $\frac{1}{M} \sum_{j \in \mathcal{J}_M(i)} Y_j(1)$  converge a  $\mu_1(X_i)$  implica que la corrección de sesgo, relativa a los estimadores por imputación, es en promedio, aproximadamente igual a  $\mu(X_i) - \hat{\mu}_1(X_i)$  lo que eliminaría cualquier inconsistencia en la regresión del estimador por imputación. En otras palabras, el estimador del emparejamiento por corrección de sesgo es robusto contra toda mala especificación de la función de regresión. Para más detalles consulte Abadie e Imbens (2011)

<sup>20</sup>Para una revisión detallada de estas propiedades se recomienda consultar Abadie e Imbens (2006), Abadie e Imbens (2011) y Abadie e Imbens (2015)

<sup>21</sup>Sin embargo, no es tan eficiente como una regresión ya que tiene la misma varianza del emparejamiento simple y tiene el problema adicional de que es difícil de implementar si la corrección de sesgo es calculada usando técnicas de suavización no paramétricas y por lo tanto requiere aproximaciones paramétricas como función del tamaño de la muestra. Pero una vez se incluyan varios polinomios e interacciones en la regresión, permitirá implementar un tipo de estimador que corregirá los problemas usuales del matching dados por la sensibilidad de estos a la elección de número de parejas y covariados

una metodología de emparejamiento genético. Adicionalmente, este documento se esfuerza no sólo corregir los problemas de especificación que soluciona un emparejamiento con corrección de sesgo, sino que también se propone usar una metodología que garantice optimizar el balance frente a otras aproximaciones cuasi-experimentales.

La idea del emparejamiento genético es mejorar un algoritmo manual presentado por Rosenbaum y Rubin (1984), quienes estipulaban que el propensity score debía ser revisado hasta que el imbalance de covariados fuera minimizado. Así mismo, a través de la literatura se muestra lo desafiante que puede llegar a ser encontrar un balance óptimo. Por ejemplo, Austin (2008) revisó 47 estudios que utilizaban emparejamiento y sólo 2 presentaban medidas de balance post-emparejamiento. Diamond y Sekhon (2005) también hicieron un análisis de investigaciones que empleaban esta metodología y de 31 artículos únicamente 11 incluían estas pruebas. Dichos autores exploran la necesidad de desarrollar un método que de manera endógena sea capaz de hallar un balance eficiente, de ahí que propongan el método de emparejamiento genético.

Básicamente, el emparejamiento genético funciona a través de un sofisticado algoritmo de búsqueda desarrollado por Mebane y Sekhon (1998) que encuentra el peso que cada covariado debería tener a la hora de estimar el propensity score. El algoritmo funciona minimizando una función generalizada de la distancia de Mahalanobis que integra un parámetro adicional de pesos,  $W$ , a saber:

$$GMD(X_i, X_j, W) = \sqrt{(X_i - X_j)^T (S^{-1/2})^T W S^{-1/2} (X_i - X_j)} \quad (9)$$

Donde  $W$ <sup>22</sup> es una matriz de pesos  $k \times k$  definida positiva,  $S$  es la matriz de covarianzas de  $X$ . Por lo tanto,  $S^{-1/2}$  es la descomposición de Cholesky de  $S$ , es decir,  $S = S^{-1/2} (S^{-1/2})^T$  y  $X^T$ . Luego, una vez son emparejados los individuos se utiliza una aproximación paramétrica que incluya los resultados de la minimización para hallar el efecto.

Teniendo claro los métodos anteriores, el conjunto de regresores usados en  $X_i$  son los reportados en las estadísticas descriptivas. El método de emparejamiento con corrección de sesgo es un método multivariado que requiere de aproximaciones paramétricas y técnicas de suavización, por lo tanto este modelo es saturado usando términos cuadráticos y al cubo para el estrato, la edad y

---

<sup>22</sup>Todos los elementos de  $W$  están restringidos a cero excepto aquellos abajo de la diagonal principal, la cual consiste de  $k$  parámetros que deben ser escogidos.

el tamaño del Hogar así como interacciones entre las variables género, área y edad. Vale la pena aclarar que estas variables además de ser referenciadas a lo largo de la literatura para explicar trabajo infantil y el acceso a otros programas, también son claves a la hora de explicar la probabilidad de recibir PAE, ya que como se detalló en la sección dedicada al funcionamiento del programa, el programa no es aleatorio sino que utiliza criterios de selección para ser asignado, por ejemplo, el programa prioriza a los niños con vulnerabilidades, de ahí que se incluya el estrato, estar en el régimen contributivo de salud, el tipo de vivienda en donde habita la familia, la presencia de los padres en el hogar entre otras. Pero lo más importante es que se incluyen variables como el área donde habita el niño, el nivel de Sisbén y el grado escolar, que finalmente si se incluyen como variables priorizadoras en los lineamientos técnicos del PAE. Así mismo se incluyen efectos fijos por municipio ( $\eta_j$ ), la inclusión de esta variable controlaría por efectos propios de los municipios, por ejemplo, los operadores o empresas encargados de preparar y suministrar el programa a los niños.

Con el ánimo de realizar comparaciones entre diferentes metodologías, en esta investigación se estimará en primer lugar el efecto a través de una regresión por MCO. Sin embargo, vale la pena aclarar que el coeficiente determinado con MCO se obtuvo después de emparejar la muestra usando el emparejamiento como una técnica de pre-procesamiento (Ho, Imai, King y Stuart, 2007).

Segundo, empleando un modelo de emparejamiento simple y finalmente usando la metodología de emparejamiento con corrección de sesgo y emparejamiento genético, para que los valores sean comparables se encontrará el efecto del tratamiento sobre los tratados (ATT) que es finalmente el hallado por el emparejamiento simple. En resumen, se encontrarán las siguientes ecuaciones<sup>23</sup>:

*Mínimos Cuadrados Ordinarios:*

$$Y_i = \alpha + \beta_1 PAE_i + \beta_2 X_i + \eta_i + \varepsilon_i \quad (10)$$

*Emparejamiento simple:*

$$\tau_{ATT}^{PSM-VC} = \sum_{i=1}^I Promedio((Y_i || D_i = 1) - (Y_{c(i)} || D_i = 0)) \quad (11)$$

---

<sup>23</sup>En el caso del emparejamiento genético después de minimizar la ecuación generalizada de la distancia de Mahalanobis se procede a encontrar el ATT utilizando la ecuación paramétrica saturada utilizando términos al cubo y al cuadrado



*Emparejamiento con corrección de sesgo:*

$$\tau_M^{bcm} = \frac{1}{N} \sum_{i=1}^N (\tilde{Y}_i(1) - \tilde{Y}_i(0) | D_i = 1) \quad (12)$$

*Emparejamiento genético:*

$$GMD(X_i, X_j, W) = \sqrt{(X_i - X_j)^T (S^{-1/2})^T W S^{-1/2} (X_i - X_j)} \quad (13)$$

## 6. Resultados

### 6.1. Impacto del Programa de Alimentación Escolar

Antes de mostrar los resultados de los modelos especificados en la sección anterior, es preciso mostrar algunos aspectos deseables en términos de soporte común, propensity score y balance de covariados <sup>24</sup>. Así, en el cuadro 4 se observa que el PAE es explicado de manera estadísticamente significativa y con el signo esperado por las variables socio económicas incluidas. En efecto, se muestra que a medida que aumenta la edad disminuye la probabilidad de tener PAE lo que es consistente con el mecanismo de asignación que favorece a los menores. En ese mismo sentido, un aumento del grado cursado disminuye la probabilidad de tener PAE; para la variable de madre en el hogar el signo también es el esperado, el hecho de que la madre viva en el hogar es una señal que puede reflejar que un niño puede tener PAE debido a que su mamá es diligente y sabe que debe inscribirlo en un colegio donde se suministre el programa; por último pertenecer al régimen contributivo y un mayor sisbén reducen la probabilidad ya que estos niños no son vulnerables.

A su vez, en la figura 5 se muestra que hay un buen soporte común, es decir, que hay una gran proporción de niños en el grupo de control y en el grupo de tratamiento con la misma probabilidad de pertenecer al programa y cuya única diferencia entre ambos grupos es recibir PAE. Además la figura 6, muestra el soporte común luego de aplicar el emparejamiento simple.

Así mismo, hay un par de discusiones interesantes relacionadas con la validez de estos resul-

---

<sup>24</sup>El soporte común se incluye en la regresión del emparejamiento con corrección de sesgo como un covariado adicional

tados. El cuadro 5 muestra, que el emparejamiento simple por vecino más cercano cierra todas las brechas en las variables observables. Es decir, que el método funciona ya que está comparando individuos más parecidos en probabilidad donde la única diferencia es recibir el PAE. En el cuadro 6 se agrega como variables explicativas el puntaje de sisbén, la educación de la madre como dummies y el área donde habita el niño (urbana o rural), y en el cuadro 7 se incluye estar en familias en acción. Que las brechas se sigan cerrando a través de las diferentes especificaciones es una señal de que el método de emparejamiento arrojará resultados creíbles y de que se cumple el supuesto tradicional de balance en variables no observables, además el porcentaje de pareo es del 90 %.

El análisis de balance entre covariados también fue realizado para el método de emparejamiento con corrección de sesgo, donde según el cuadro 8 permanece lo expuesto por Abadie e Imbens (2011), quienes encuentran que la calidad del emparejamiento va en dirección opuesta al número de parejas. Se observa por ejemplo, que las brechas a niveles de covariados se cierran en su totalidad para los números más pequeños de parejas y a medida que  $m$  crece, diferencias de medias vuelven a ser significativas. Vale la pena mencionar que Abadie e Imbens repetidamente prefieren utilizar cuatro parejas, ya que dicho ajuste ofrece la ventaja de no tener que depender de muy poca información sin incorporar las observaciones que no son lo suficientemente similares (Abadie e Imbens 2006 y 2011).

En cuanto al emparejamiento genético, se observa que el método cierra todas las brechas después de emparejar la muestra. Finalmente, la variable con la brecha más grande después del emparejamiento fue vivienda propia, sin embargo, esta no es estadísticamente significativa<sup>25</sup>.

Luego de que ya se presentó evidencia que hace más plausible la validez de estos resultados, se considera conveniente mostrarlos. Por lo tanto, se presenta el cuadro 9, este contiene cada uno de los resultados para trabajo infantil usando toda la muestra y a través de cada fila se observan los resultados provenientes de cada método empleado partiendo desde MCO hasta el emparejamiento genético<sup>26</sup>. En ese orden de ideas encontrar un resultado parecido y dentro del rango de los coeficientes hallados en esta investigación para los métodos de emparejamiento con reducción de sesgo y emparejamiento genético, es otra muestra de que el resultado es robusto. Ahora, ya que es este efectivamente un valor cercano, finalmente se puede concluir que el PAE reduce el trabajo infantil

---

<sup>25</sup> Antes la brecha de esta variable era significativa con un p-value del 0.002, después del emparejamiento el p-value arrojado fue de 0.488

<sup>26</sup> Tanto el emparejamiento simple Vecino más Cercano como el de kernell utilizan todas las variables usadas del cuadro 5 al 7 y efectos fijos de municipio

alrededor del 4 %.

Se aprecia que aunque el resultado por MCO arroja un coeficiente del 5 % este a su vez va disminuyendo y se ubica robustamente en un casi 4 % a través de los distintos métodos de emparejamiento de mayor calidad. Es de notar que dado que la bondad del emparejamiento con reducción de sesgo disminuye con el número de parejas empleadas, se prefiere usar el estimador producto de los 4 vecinos, es decir, un coeficiente de -3,7 %, lo que es bastante cercano al estimador del emparejamiento genético.

Estos resultados ponen de manifiesto que efectivamente existía el problema de endogeneidad descrito para el método de MCO, que como se suponía, sobrestimaba el coeficiente incluso en mayor magnitud que el emparejamiento simple. Así mismo, en cuanto al emparejamiento simple, encontrar un resultado mediante emparejamiento con corrección de sesgo y emparejamiento genético que es menor a este, muestra que efectivamente a pesar de que las condiciones sean favorables para emplear la metodología, en cuanto a amplio soporte común y cierre de brechas, no se puede dejar de lado el sesgo que implica usarlo (Smith y Todd, 2005). Sin embargo, gracias a las variantes de estimadores que se utilizaron en esta investigación, se procedió a encontrar un robusto coeficiente cercano al 4 % lo que significaría que efectivamente el PAE reduce el trabajo infantil en esta magnitud.

Si bien este coeficiente parece modesto, es superior en términos de magnitud y significancia para otros trabajos que han investigado programas tipo PAE y no encuentran resultados estadísticamente significativos (Jacoby, 2002; Afridi, 2010 y McEwan, 2013). Ahora, en el caso de Familias en Acción, en Colombia, el coeficiente es mayor que lo reportado por Attanasio y Gómez (2004) para el caso de familias en acción y su impacto en el trabajo infantil, que no encuentra ningún efecto. Inclusive, el resultado es mayor que lo reportado por Quiroga (2006) para los niños que sólo trabajan (2 %).

Ahora bien, partiendo del marco que el Programa de Alimentación Escolar efectivamente sí reduce el trabajo infantil, es preciso encontrar poblaciones en las que se potencializa este efecto. Para encontrar estos grupos, se procedió a fraccionar la muestra entre hombres y mujeres, entre áreas rural y urbana, entre secundaria y primaria y finalmente para los niños que viven en el área rural y cursan algún grado de secundaria. Una vez, la muestra fue dividida se aplicó la correspon-

diente metodología<sup>27</sup>. Los resultados de este procedimiento se presentan en el cuadro 10, cuyas filas corresponden al método de estimación y las columnas son las muestras que se emplearon. Precisamente, se encuentran efectos mayores cuando los grupos son más vulnerables y expuestos al trabajo infantil, como es ser hombre, habitar el área rural y estar en secundaria. Así se encuentra que el programa reduce el trabajo infantil en los hombres en un 7%, en el área rural en un 9%, en secundaria en un 6% y finalmente en los niños que cursan secundaria que viven en área rural el efecto es cercano al 11%, resultados que difieren de los encontrados con regresión y emparejamiento simple que respectivamente son 8%, 13%, 7% y 16%<sup>28</sup>. Estos coeficientes superan en términos de magnitud y significancia a lo hallado por Attanasio y Gómez (2004), quienes usan efectos heterogéneos para área y género para el programa Familias en Acción.

## 6.2. Mecanismos

Los resultados anteriores encajan en el marco teórico expuesto en secciones anteriores. Que el programa reduzca la participación laboral no sólo indica que se cumplen los objetivos que buscan mejorar la vida académica de los niños y las niñas, ya que los aleja del trabajo infantil y pueden tener más tiempo para dedicarse a actividades que les permitan desarrollar sus capacidades físicas y socio emocionales. También es muestra de que el PAE en el caso de Colombia, tiene efectos importantes en el contexto económico de las familias. En la medida en que los hogares con niños elegidos para PAE, que por definición son vulnerables, puedan acceder a un desayuno o almuerzo gratis, esto influye en la toma de las decisiones de las familias, ya que este tipo de transferencia en especie cambia la riqueza de los hogares, generando que estas unidades se sientan menos pobres y con la capacidad de enviar al niño al colegio en vez de al mercado laboral.

Por lo tanto, dicho cambio en las decisiones de los hogares debe manifestarse no sólo en la reducción del trabajo infantil, sino también en un correspondiente aumento de la asistencia escolar de los niños y las niñas así como de actividades que les permitan desarrollar sus capacidades físicas. Con el ánimo de contrastar este mecanismo se presentan los resultados para la variable dependiente asistencia escolar en el cuadro 11, columna 1, y en cuanto a variables que involucran actividad física, se presentan resultados desde la columna 3 a la 5. En estas columnas se muestra

---

<sup>27</sup>Dado el diseño representativo de la encuesta a nivel nacional, las submuestras siguen siendo representativas (ENDS,2010)

<sup>28</sup>El soporte común correspondiente para cada muestra empleada es mostrado en el Apéndice 2.

el aumento esperado de variables que determinan el desarrollo físico y emocional de los niños y las niñas, por lo tanto, es de notar que el PAE tiene un efecto deseable ya que una disminución del trabajo infantil se complementa con un correspondiente aumento en actividades escolares y extracurriculares de los niños.

Ahora bien, ya que esta investigación plantea que son los niños pobres los que presentan mayor participación laboral infantil a nivel rural y urbano, este argumento va en contra de lo expuesto por Bhalotra y Heady (2003), quienes muestran para Pakistán y Ghana que las niñas en hogares con mayor acumulación de tierra, tienden a trabajar más, lo que sugeriría que la reducción del trabajo infantil es debida a diferencias en la riqueza no observable y no al programa. Por lo tanto, aunque ya se pudo establecer que en Colombia los niños pobres trabajan más, lo correcto es mostrar como el PAE incrementa la riqueza de los hogares, y como se hizo mención antes dicho resultado se mostrará haciendo uso de una variable que mide la seguridad alimentaria de los hogares. Los resultados para la seguridad alimentaria de los hogares son presentados a su vez en la segunda columna del cuadro 11, donde se observa que efectivamente el Programa de Alimentación Escolar mantiene una relación causal con la seguridad alimentaria nutricional aumentando esta percepción en hogares con PAE.

Que el programa disminuya el trabajo infantil y aumente la asistencia, es un indicio de que las familias se sienten más seguras y son capaces de realizar pequeñas inversiones en sus hijos como es permitirlos asistir al colegio. Luego, ya que el programa de alimentación escolar es una estrategia de tipo nutricional se esperaría que mejorara las características alimenticias de la familia. Ya que a uno o varios de los miembros del hogar se les suministre una comida diaria, las unidades familiares podrían ahorrar algo de dinero o dedicar esta comida a algún otro integrante, pero dicha disponibilidad del alimento es precisamente señal de seguridad alimentaria. Por lo tanto, dicho análisis concluiría que el hogar se considera más seguro en términos nutricionales, y que mejor variable para medir esta percepción que la variable de seguridad alimentaria.

Vale la pena aclarar, que la variable de seguridad alimentaria también fue considerada como una variable mediadora. Es decir, el mecanismo a través del cual el Programa de Alimentación Escolar reduce el trabajo Infantil. Para determinar que efectivamente esta recogía el impacto del PAE fue incluida junto con las otras variables socio económicas. Luego, se encontró que el coeficiente pasó de ser estadísticamente significativo al 1 % a estar al 10 % en todas las metodologías. Por lo tanto, se puede concluir que uno de los canales que el PAE utiliza para reducir el trabajo infantil es a

través de aumentar la seguridad alimentaria. Sin embargo, aunque uno de los canales en que el PAE actúa para disminuir el trabajo infantil es la seguridad alimentaria, este análisis también da a lugar a que existen mecanismos alternativos en que actúa el programa, como puede ser la nutrición de los niños, aspecto no explorado en esta investigación dado que pertenece al enfoque tradicional del PAE, ya antes discutido.

Ahora, que el programa reduzca el trabajo infantil y aumente la asistencia, la seguridad alimentaria de los hogares y la actividad física de los niños es una muestra de los beneficios a corto plazo del PAE. Por lo tanto, sí se piensa en un análisis costo-beneficio de la intervención, se debe ir más allá del costo cercano al billón de pesos mensuales que el programa costaba para la época y más bien, se deben cuantificar los beneficios a largo plazo que trae tanto para los niños y sus familias e inclusive las externalidades positivas generadas para la sociedad entera (CINDE, 2013).

### **6.3. Pruebas de Robustez**

Finalmente, a manera de realizar pruebas de robustez se presenta una prueba gráfica de si se cumple el supuesto de que al controlar por características observables se logra balancear las no observables, es preciso observar lo que ocurre en las figuras 7 y 8, donde se observa que una vez se utiliza el método, las distribuciones de variables para controles y tratados no incluidas en la especificación, como es el caso de la Talla en centímetros de los niños y el índice de masa corporal, parecen también haberse aproximado y mostrar un mejor balance. Además, para estas mismas variables se realiza una diferencia de medias, donde se comprueba que efectivamente se cerraron las brechas en controles que incluso no fueron incluidos en las especificación, hecho que se comprueba en el cuadro 12.

De igual manera, es importante realizar algún tipo de prueba de falsificación, en esta oportunidad se usará una variable de resultado ficticia, es decir, una prueba de pseudoresultado. Imbens y Rubin (2010), muestran que para realizar dicha prueba se debe estimar el efecto del programa sobre una variable de resultado ficticia, de la cual tenemos la seguridad de que no puede ser afectada por el tratamiento y este, a su vez, no debe ser significativo. En esta ocasión se utilizó una variable indicadora que tiene en cuenta si los niños sufren de alguna discapacidad y se procedió a incluir también controles que sí determinen esta condición, como si la madre fumaba o tomaba, el estado de salud actual, el peso de las madres en relación con la edad, entre otras. Tal como se muestra en

el cuadro 13, el tratamiento no tiene ningún efecto en la probabilidad de que el niño o la niña tenga una discapacidad, por lo que el método de emparejamiento simple y con reducción de sesgo son adecuados en este caso.

Finalmente, los cuadros 14 al 17 muestran pruebas de distribución de kolmogorov-smirnov. El cuadro 14 muestra los p-values de las variables incluidas en el modelo de emparejamiento simple, donde se observa que antes del emparejamiento se rechazaba la hipótesis nula de que las distribuciones eran iguales entre grupos, mientras que después del emparejamiento ésta no se rechaza. Así mismo, en el cuadro 15 se aprecia que tal como se cierran las brechas en medias para variables que no se tienen en cuenta en el propensity score, también sucede para su distribución.

Por su parte, en el cuadro 16 se encuentra lo mismo para el emparejamiento con corrección de sesgo, de nuevo notando que la calidad del emparejamiento empeora a medida que aumenta el número de vecinos. Por último, el cuadro 17 muestra esta prueba para el emparejamiento genético, donde se aprecia que el algoritmo usado para hallar los pesos que toman los covariados también fue óptimo, esto debido a que las distribuciones son iguales para tratados y controles después del emparejamiento.

## **7. Conclusiones**

Teniendo en cuenta los resultados anteriores así como el marco teórico expuesto y la revisión de literatura, se realizarán ciertas discusiones imprescindibles para concluir esta investigación. En primer lugar, partiendo de aquellos estudios que proponen cambiar la manera en que se piensan y evalúan los Programas de Alimentación Escolar, se demostró que el PAE reduce el trabajo infantil, mejorando la percepción de seguridad alimentaria de los hogares. Dicho resultado, también se relaciona con la revisión de literatura propuesta, ésta señalaba cómo el desarrollo físico de los niños y las niñas afecta su parte psico-emocional. De tal manera que, sí el alimento entregado por el Programa de Alimentación Escolar que tiene implicaciones físicas, es capaz de cambiar el comportamiento o percepción de las familias, esto es la seguridad alimentaria, la teoría se mantiene para este documento, porque las decisiones de los hogares permitirán que el niño o la niña siga asistiendo al colegio en vez de enviarlo al mercado laboral y asegurarle unas precarias condiciones en su vida futura.

Así mismo, se hace necesario hablar de la metodología empleada, en este caso ya que este documento ofreció comparaciones entre regresión, emparejamiento simple, emparejamiento con corrección de sesgo y emparejamiento genético, es importante mencionar que además de las ventajas explicadas en la sección correspondiente a estrategia empírica, los dos últimos estimadores se comportan de manera deseable en términos de balance de covariados. Al mismo tiempo, ambos reporta coeficientes más bajos, pasando de un coeficiente obtenido por regresión del 5,1 % y uno de emparejamiento simple del 5 % a finalmente un coeficiente que siempre es cercano al 4 % con emparejamiento con reducción del sesgo y emparejamiento genético. Que el coeficiente varíe entre métodos, pero que se mantenga entre diferente número de vecinos para el caso del emparejamiento con corrección de sesgo, demuestra que la disminución del trabajo infantil vía PAE es robusta en magnitud y signo.

Finalmente, es importante abordar la discusión de por qué otras investigaciones acerca del Programa de Alimentación Escolar no encuentran resultados significativos. Como se mencionó en apartados anteriores, muchos de los estudios relacionados con programas similares no tienen en cuenta la multidimensionalidad del PAE. Por ejemplo, sería incorrecto evaluar el efecto del PAE en variables físicas y sociales sin considerar todos los mecanismos que se llevan a cabo en la familia. Probablemente, si los estudios considerarán que estos programas implican una redistribución de alimentos hacia otros miembros del hogar como sus hermanos menores, estos sí hallarán efectos significativos. Esta investigación, por su parte reconoce el costo de oportunidad que enfrentan las familias a la hora de enviar a los niños y niñas al colegio, que es reducido mediante el PAE y que se refleja en una disminución estadísticamente significativa del trabajo infantil.

Que un programa sea capaz de combinar herramientas que mejoren los aspectos físicos, emocionales, mentales y sociales de los niños, niñas y adolescentes significa que es una poderosa herramienta, que va más allá de las típicas preguntas que incluyen los estudios cuando se trata de investigar programas con un complemento nutricional. Lo anterior, sugiere que es importante entender primero la metodología y campo de acción de un programa de manera amplia, para así comprender finalmente tanto los efectos convencionales como no convencionales, impactos que aunque no se deduzcan inmediatamente no quiere decir que no sean aportes significativos en favor de los más y las más vulnerables.



## 8. Cuadros y Gráficas

Cuadro 1: Estadísticas Descriptivas

Variable	Promedio	Des. Est.	Min	Max
Variable Dependiente				
Trabajo infantil	0.086	0.28	0	1
Tratamiento				
PAE	0.52	0.5	0	1
Variables Socio Demográficas				
Edad	11.039	3.707	5	17
Hombre	0.507	0.5	0	1
Último grado cursado	4.977	3.111	0	11
Madre en el hogar	0.834	0.372	0	1
Padre en el hogar	0.616	0.486	0	1
Educación de la madre	1.686	0.72	0	3
Tamaño en el hogar	5.639	2.32	1	21
Número de Adultos mayores en el hogar	0.257	0.55	0	4
Vivienda Propia	0.615	0.487	0	1
Estrato	1.647	0.871	0	6
Régimen Contributivo	0.301	0.459	0	1
Asistencia	0.92	0.26	0	1
Área Urbana	0.66	0.47	0	1
Secundaria	0.56	0.49	0	1
Percibe Ingresos	0.66	0.47	0	1
Sisbén	1.96	1.24	1	4

Número de observaciones 47,583

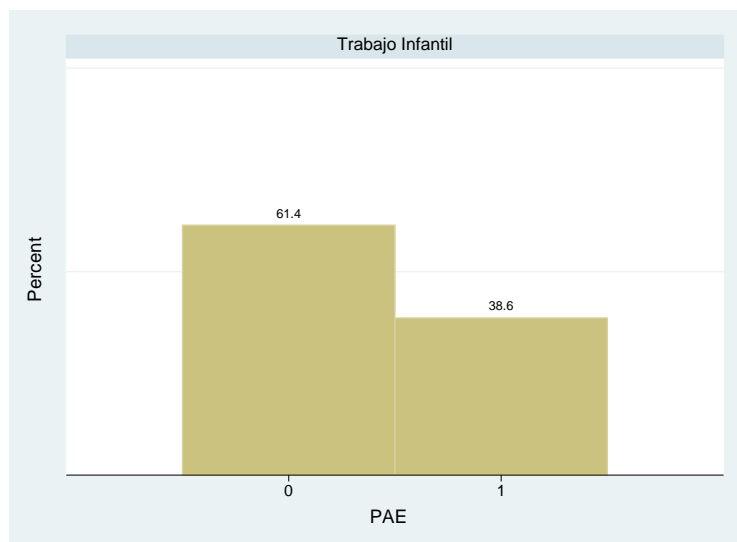
Las variables de madre y padre en el hogar son indicadoras iguales a 1 si la familia cuenta con estos miembros. La variable régimen contributivo es igual a 1 si la persona pertenece al régimen contributivo mientras que asistencia y secundaria son variables dummies igual a 1 si el niño asiste al colegio y está en secundaria, respectivamente. Por otro lado, la variable percibe ingresos toma el valor de 1 si la persona recibe ingresos periódicamente. Además, la variable Educación de la madre va de 0 a 3, donde 0 es ninguna educación, 1 es primaria, 2 es secundaria y 3 es educación universitaria. Así mismo la variable tamaño del hogar mide el número total de miembros que viven en la casa, esto también aplica para los adultos mayores de 65 años en el hogar. Finalmente, la variable sisbén comprende los niveles de 1 hasta 4.

Cuadro 2: Diferencias de medias

Variable	Prom. Sin PAE	Prom. con PAE	Diferencia
Edad	11.991 (0.024)	10.162 (0.021)	1.965*** (0.319)
Hombre	0.503 (0.003)	0.5106 (0.003)	-0.007* (0.004)
Ultimo grado cursado	5.708 (0.021)	4.240 (0.020)	1.46*** (0.029)
Madre en el Hogar	0.817 (0.002)	0.849 (0.002)	-0.031*** (0.003)
Padre en el Hogar	0.597 (0.003)	0.633 (0.002)	-0.036*** (0.004)
Tamaño del Hogar	5.409 (0.014)	5.871 (0.014)	-0.461*** (0.02)
Adulto Mayor en el Hogar	0.280 (0.003)	0.237 (0.003)	0.004*** (0.004)
Vivienda Propia	0.600 (0.003)	0.631 (0.002)	-0.030*** (0.004)
Percibe Ingresos	0.715 (0.004)	0.469 (0.005)	0.246*** (0.006)
Estrato	1.886 (0.006)	1.403 (0.004)	0.718*** (0.004)
Régimen contributivo	0.422 (0.003)	0.182 (0.002)	0.240*** (0.003)
Sisben	2.882 (0.008)	1.752 (0.006)	0.430*** (0.010)
Madre con primaria	0.253 (0.003)	0.442 (0.003)	-0.189*** (0.002)
Madre con secundaria	0.548 (0.003)	0.448 (0.003)	0.099*** (0.002)
Madre con universidad	0.169 (0.002)	0.060 (0.001)	0.108*** (0.002)
Asistencia	0.850 (0.002)	0.991 (0.000)	-0.140*** (0.002)
Área Urbana	0.755 (0.002)	0.559 (0.003)	-0.220*** (0.002)
Trabajo Infantil	0.108 (0.002)	0.064 (0.001)	0.043*** (0.002)

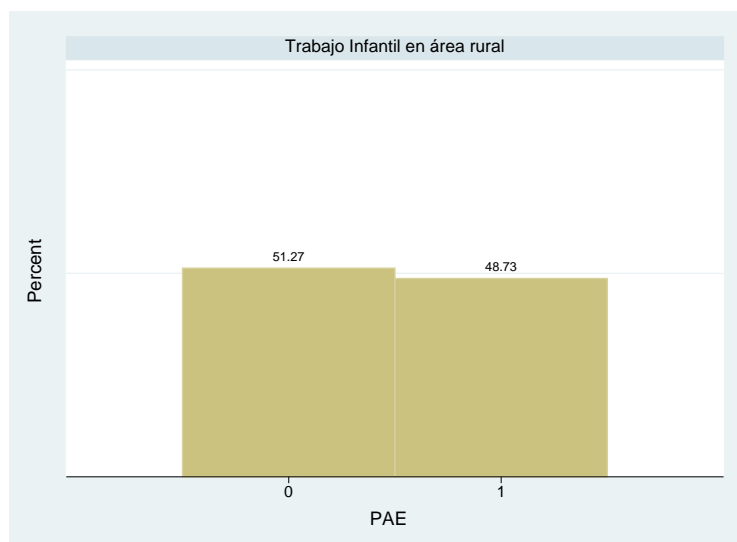
Error Estándar en paréntesis . \* representa significancia estadística al 10 %, \*\* al 5 % y \*\*\* al 1 %.

Figura 1: Niños y niñas que trabajan sin PAE vs Niños y niñas que trabajan con PAE



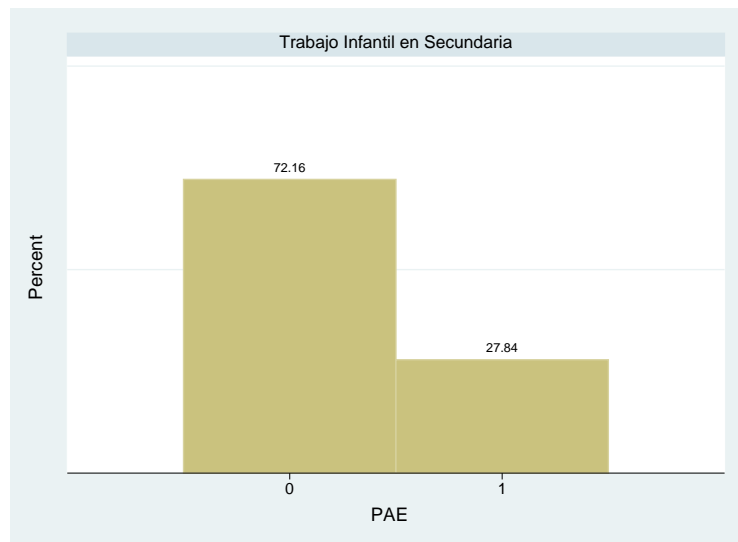
Donde 0 es para los niños que no tienen PAE y 1 para los que si tienen PAE

Figura 2: Niños y niñas que trabajan sin PAE vs Niños y niñas que trabajan con PAE en el sector rural



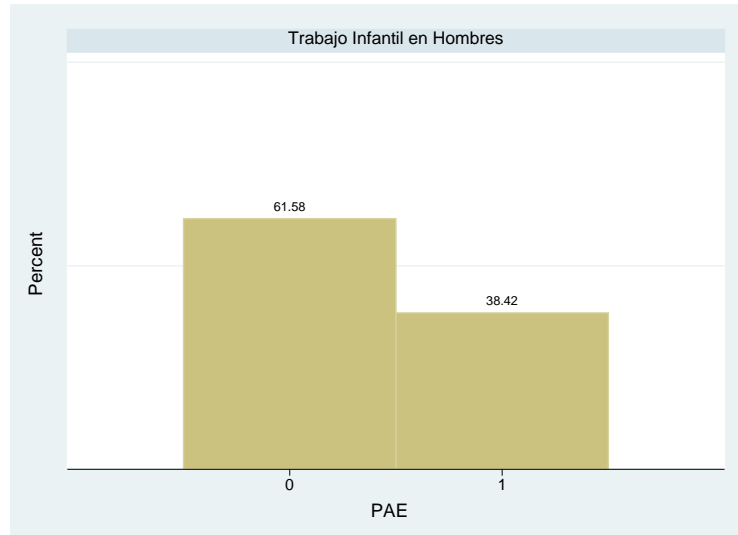
Donde 0 es para los niños que no tienen PAE y 1 para los que si tienen PAE

Figura 3: Niños y niñas que trabajan sin PAE vs niños y niñas que trabajan con PAE en Secundaria



Donde 0 es para los niños que no tienen PAE y 1 para los que si tienen PAE

Figura 4: Niños que trabajan sin PAE vs Niños que trabajan con PAE (Hombres)



Donde 0 es para los niños que no tienen PAE y 1 para los que si tienen PAE

Cuadro 3: Estadísticas descriptivas de la Seguridad Alimentaria

Variable	Seguridad alimentaria
	Área
Rural	2.71
Urbano	3.09
Programa de Alimentación Escolar	
Sin PAE	3.03
Con PAE	2.71
Trabajo Infantil	
No trabaja	2.87
Trabaja	2.06
Trabajo Infantil en el área rural	
No trabaja	2.66
Trabaja	2.46
Trabajo Infantil en niños con PAE	
Sin PAE	2.72
Con PAE	2.44

Se muestran los promedios de la variable seguridad alimentaria, una escala que va de 1 a 4 donde 1 es el nivel de inseguridad alimentaria más severo y 4 es que el hogar presenta seguridad alimentaria.

Cuadro 4: Propensity Score

Variable	Variable dependiente PAE
Edad	-0.211*** (-0.008)
Hombre	-0.007 (0.027)
Ultimo Grado Cursado	-0.027*** (0.008)
Madre en el Hogar	0.475*** (0.062)
Padre en el Hogar	0.035 (0.026)
Tamaño del Hogar	0.020*** (0.006)
Adulto Mayor en el Hogar	-0.127*** (0.026)
Vivienda Propia	-0.016 (0.026)***
Estrato	-0.43*** (0.017)
Régimen contributivo	-0.655** (0.030)
Sisben	-0.090*** (0.011)
Madre con Primaria	0.303*** (0.076)
Madre con Secundaria	0.138** (0.078)
Madre con Superior	-0.486*** (0.088)
Area	-0.611** (0.029)

Error Estándar en paréntesis . \* representa significancia estadística al 10 %, \*\* al 5 % y \*\*\* al 1 %.

Figura 5: Propensity Score trabajo Infantil

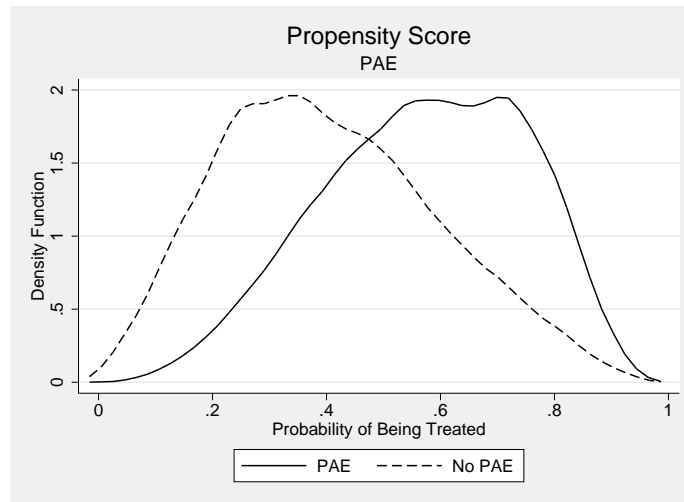
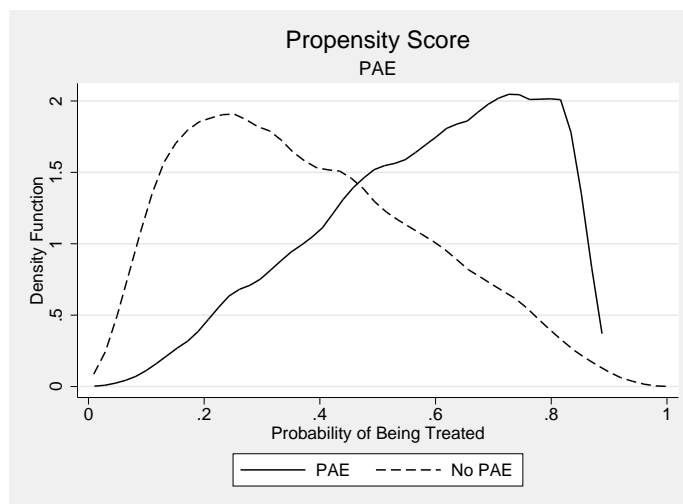


Figura 6: Propensity Score trabajo Infantil después del matching



Se borran las observaciones que no están en el soporte común mediante la función `trimming(10)`

Cuadro 5: Diferencia de medias entre tratados y controles antes y después del emparejamiento simple

Variabes	Sin Matching (U)/Con Matching (M)	Tratados	Control	Sesgo	P-Value
Edad	U	10.160	11.991	-50.8	0.000
	M	11.788	11.819	-0.9	0.408
Hombre	U	0.511	0.504	1.4	0.012
	M	0.484	0.491	-1.5	0.237
Último grado	U	4.242	5.708	-48.5	0.000
	M	4.759	4.750	0.3	0.784
Madre en el Hogar	U	0.849	0.81832	8.4	0.000
	M	0.964	0.962	0.6	0.357
Padre en el Hogar	U	0.634	0.596	7.7	0.000
	M	0.653	0.652	0.3	0.810
Educación de la madre	U	1.523	1.863	-48.5	0.000
	M	1.649	1.648	0.3	0.832
Tamaño en el hogar	U	5.864	5.395	20.4	0.000
	M	5.661	5.688	-1.1	0.364
Adulto Mayor en el hogar	U	0.236	0.279	-7.8	0.000
	M	0.204	0.206	-0.4	0.714
Vivienda Propia	U	0.629	0.598	6.4	0.000
	M	0.596	0.592	0.8	0.541
Estrato	U	1.407	1.893	-58.0	0.000
	M	1.511	1.496	1.8	0.100
Régimen contributivo	U	0.184	0.428	-54.7	0.000
	M	0.257	0.254	0.7	0.563



Cuadro 6: Diferencia de medias entre tratados y controles antes y después del emparejamiento.  
*Sisbén, Área Urbana y Educación de la madre*

Variablen	Sin Matching (U) /Con Matching(M)	Tratados	Control	Sesgo	P-Value
Edad	U	10.161	11.997	-5.8	0.000
	M	11.799	11.820	-0.9	0.585
Hombre	U	0.511	0.503	1.4	0.096
	M	0.481	0.490	-1.5	0.176
Último grado	U	4.241	5.708	-48.5	0.000
	M	4.768	4.747	0.3	0784
Madre en el Hogar	U	0.849	0.817	8.4	0.000
	M	0.964	0.962	0.6	0.377
Padre en el Hogar	U	0.634	0.597	7.7	0.000
	M	0.656	0.653	3.3	0.666
Tamaño del hogar	U	5.871	5.409	20.4	0.000
	M	5.677	5.703	-1.1	0.348
Vivienda Propia	U	0.631	0.601	6.4	0.000
	M	0.597	0.594	0.8	0.621
Estrato	U	1.404	1.886	-5.8	0.000
	M	1.503	1.490	1.8	0.141
Régimen contributivo	U	0.182	0.423	-7.9	0.000
	M	0.255	0.252	1.8	0.524
Adulto Mayor en el hogar	U	0.236	0.279	-7.8	0.000
	M	0.198	0.203	-0.4	0.601
Madre en primaria	U	0.443	0.253	40.7	0.000
	M	0.371	0.370	0.2	0.861
Madre en secundaria	U	0.449	0.549	-20.1	0.000
	M	0.526	0.523	0.6	0.637
Madre en Superior	U	0.061	0.169	-34.6	0.000
	M	0.072	0.074	-0.7	0.469
Sisbén	U	1.752	2.182	-35.3	0.000
	M	1.940	1.926	0.6	0601
Área Urbana	U	0.553	0.764	-45.9	0.000
	M	0.687	0.685	0.5	0.724

Cuadro 7: Diferencia de medias entre tratados y no tratados antes y después del emparejamiento.  
*Incluyendo Familias en Acción*

Variablen	Sin matching (U)/Con Matching (M)	Tratados	Control	Sesgo	P-Value
Edad	U	10.16	11.99	-5.8	0.000
	M	11.429	11.439	-0.9	0.815
Hombre	U	0.510	0.503	1.4	0.120
	M	0.492	0.498	-1.5	0.408
Último grado	U	4.241	5.707	-48.5	0.000
	M	44.226	4.406	0.3	0.670
Madre en el Hogar	U	0.849	0.818	8.4	0.000
	M	0.971	0.968	0.6	0.309
Padre en el Hogar	U	0.634	0.596	7.7	0.000
	M	0.654	0.651	3.3	0.576
Educación de la madre	U	1.523	1.863	-48.5	0.000
	M	1.649	1.648	0.3	0.389
Tamaño en el hogar	U	5.865	5.395	20.1	0.000
	M	5.610	5.598	-1.0	0.705
Adulto Mayor en el hogar	U	0.236	0.279	-7.8	0.000
	M	0.204	0.206	-0.4	0.714
Vivienda Propia	U	0.631	0.601	6.4	0.000
	M	0.597	0.594	0.8	0.621
Estrato	U	1.404	1.886	-5.8	0.000
	M	1.503	1.490	1.8	0.141
Régimen contributivo	U	0.182	0.423	-7.9	0.000
	M	0.255	0.252	1.8	0.524
Familias en Acción	U	0.230	0.279	7.5	0.000
	M	0.197	0.203	0.5	0.441

Cuadro 8: Diferencias de medias después del emparejamiento con corrección de sesgo

Variable	m=4	m=16	m=64	m=256
Grado Escolar	0.004 (0.005)	0.011*** (0.002)	0.018*** (0.003)	0.019*** (0.050)
Edad*área	0.008 (0.028)	0.024* (0.022)	0.111*** (0.011)	0.115*** (0.019)
Área*género	0.001 (0.002)	0.001 (0.002)	0.003 (0.011)	0.004 (0.014)
Madre en el Hogar	0.000 (0.002)	0.000 (0.001)	0.030** (0.002)	0.040*** (0.019)
Padre en el Hogar	0.001 (0.002)	0.019 (0.011)	0.022* (0.015)	0.024*** (0.012)
Educación de la madre	0.002 (0.003)	0.027*** (0.002)	0.040*** (0.003)	0.043*** (0.005)
Tamaño del Hogar	-0.002 (0.004)	-0.212*** (0.008)	-0.215*** (0.008)	-0.223*** (0.009)
Adulto Mayor en el hogar	-0.002 (0.003)	-0.042*** (0.003)	-0.053*** (0.003)	-0.054*** (0.025)
Vivienda Propia	0.000 (0.001)	0.035*** (0.001)	0.089*** (0.005)	0.093*** (0.006)
Estrato	0.009 (0.011)	0.012 (0.027)	0.021 (0.029)	0.067*** (0.035)
Régimen Contributivo	0.000 (0.000)	0.012*** (0.000)	0.033*** (0.000)	0.041*** (0.011)
Edad	0.000 (0.006)	0.008* (0.007)	0.011*** (0.007)	0.015*** (0.009)
Edad <sup>2</sup>	0.345 (1.123)	1.589* (1.278)	3.416*** (1.127)	3.625*** (1.777)
Edad <sup>3</sup>	1.245 (2.301)	2.625* (2.005)	6.244*** (2.295)	6.637*** (2.121)
Tamaño Hogar <sup>3</sup>	1.241 (1.211)	3.841*** (1.707)	4.125*** (1.736)	6.637*** (1.121)
Estrato <sup>2</sup>	-0.023 (0.034)	-0.090 (0.101)	-0.412 (0.701)	-0.425 (0.679)
Estrato <sup>3</sup>	0.001 (0.002)	0.176*** (0.029)	0.217*** (0.004)	0.368*** (0.012)
Área	0.000 (0.001)	0.012 (0.013)	0.028*** (0.001)	0.030*** (0.000)
Agua	0.001 (0.002)	0.012*** (0.001)	0.044*** (0.001)	0.056*** (0.001)
Etnia	0.000 (0.001)	0.016*** (0.001)	0.018*** (0.000)	0.161*** (0.005)
Género	0.000 (0.002)	0.080*** (0.012)	0.092*** (0.028)	0.099*** (0.000)
Soporte Común	0.001 (0.002)	0.070*** (0.012)	0.099*** (0.015)	0.099*** (0.001)

En la tabla se muestran las diferencias de medias para cada vecino m. Error Estándar en paréntesis . \* representa significancia estadística al 10%, \*\* al 5% y \*\*\* al 1%. La m corresponde al número de vecinos empleados.

Cuadro 9: Efecto del Programa de Alimentación Escolar en el Trabajo Infantil

Método	Impacto en Trabajo Infantil
Mínimos Cuadrados Ordinarios	-0.051*** (0.003)
Emparejamiento simple con VMC	-0.049*** (0.003)
Emparejamiento simple con kernell	-0.053*** (0.004)
Emparejamiento con reducción de sesgo m=1	-0.037*** (0.004)
Emparejamiento con reducción de sesgo m=4	-0.037*** (0.003)
Emparejamiento con reducción de sesgo m=16	-0.038*** (0.002)
Emparejamiento con reducción de sesgo m=64	-0.040*** (0.003)
Emparejamiento con reducción de sesgo m=256	-0.040*** (0.003)
Emparejamiento genético	-0.039*** (0.001)

Error Estándar en paréntesis . \* representa significancia estadística al 10 %, \*\* al 5 % y \*\*\* al 1 %. VMC es el emparejamiento por vecino más cercano y m es el número de vecinos empleados

Cuadro 10: Efectos heterogéneos del Programa de Alimentación Escolar en el Trabajo Infantil

Método	Hombres	Rural	Secundaria	Rural y Secundaria
Mínimos Cuadrados Ordinarios	-0.082*** (0.004)	-0.129*** (0.006)	-0.069*** (0.004)	-0.165*** (0.010)
Emparejamiento simple con VMC	-0.078*** (0.006)	-0.122*** (0.011)	-0.063*** (0.006)	-0.155*** (0.015)
Emparejamiento simple con kernell	-0.071*** (0.001)	-0.127*** (0.010)	-0.068*** (0.001)	-0.150*** (0.000)
Emparejamiento con reducción de sesgo m=1	-0.066*** (0.006)	-0.094*** (0.011)	-0.064*** (0.005)	-0.115*** (0.014)
Emparejamiento con reducción de sesgo m=4	-0.065*** (0.005)	-0.093*** (0.009)	-0.062*** (0.005)	-0.103*** (0.012)
Emparejamiento con reducción de sesgo m=16	-0.070*** (0.004)	-0.092*** (0.010)	-0.064*** (0.004)	-0.104*** (0.018)
Emparejamiento con reducción de sesgo m=64	-0.074*** (0.003)	-0.093*** (0.011)	-0.065*** (0.004)	-0.103*** (0.022)
Emparejamiento con reducción de sesgo m=256	-0.075*** (0.003)	-0.094*** (0.012)	-0.065*** (0.004)	-0.102*** (0.022)
Emparejamiento genético	-0.079*** (0.002)	-0.101*** (0.001)	-0.060*** (0.003)	-0.151*** (0.001)

Error Estándar en paréntesis . \* representa significancia estadística al 10%, \*\* al 5% y \*\*\* al 1%. VMC es el emparejamiento por vecino más cercano y m es el número de vecinos empleados

Cuadro 11: Mecanismos del PAE a través de la Asistencia Escolar y seguridad alimentaria

Modelo	Asistencia Escolar	Seguridad Alimentaria	Caminó en tiempo libre	Practica actividad física moderada	Practica actividad física rigurosa
Mínimos cuadrados ordinarios	0.174*** (0.006)	0.061*** (0.001)	0.181*** (0.003)	0.047*** (0.001)	0.065*** (0.006)
Emparejamiento simple kernell	0.157*** (0.004)	0.044*** (0.001)	0.163*** (0.003)	0.038*** (0.002)	0.057*** (0.007)
Emparejamiento simple VMC	0.144*** (0.005)	0.034*** (0.001)	0.161*** (0.004)	0.037*** (0.003)	0.051*** (0.007)
Emparejamiento con corrección de sesgo m=1	0.131*** (0.005)	0.023*** (0.001)	0.162*** (0.003)	0.034*** (0.002)	0.055*** (0.005)
Emparejamiento con corrección de sesgo m=4	0.132*** (0.003)	0.023*** (0.003)	0.162*** (0.004)	0.033*** (0.004)	0.049*** (0.003)
Emparejamiento con corrección de sesgo m=64	0.145*** (0.041)	0.025*** (0.001)	0.172*** (0.002)	0.033*** (0.005)	0.049*** (0.006)
Emparejamiento con corrección de sesgo m=256	0.165*** (0.039)	0.025*** (0.003)	0.175*** (0.002)	0.037*** (0.005)	0.065*** (0.007)
Emparejamiento genético	0.135*** (0.005)	0.023*** (0.001)	0.162*** (0.004)	0.039*** (0.004)	0.049*** (0.005)

Error Estándar en paréntesis . \* representa significancia estadística al 10 %, \*\* al 5 % y \*\*\* al 1 %. VMC es el emparejamiento por vecino más cercano y m es el número de vecinos empleados. La columna (1) corresponde a los resultados estimados para seguridad alimentaria mientras la segunda (2) es para asistencia escolar. La tercera (3) muestra los resultados estimados correspondientes a sí el niño caminó en su tiempo libre, la cuarta (4) es una variable dummy que es igual a 1 sí el niño practicó actividad física moderada la semana pasada a la encuesta y la quinta (5) toma el valor de 1 sí el niño practicó actividad física rigurosa

Figura 7: Distribución antes y después del emparejamiento

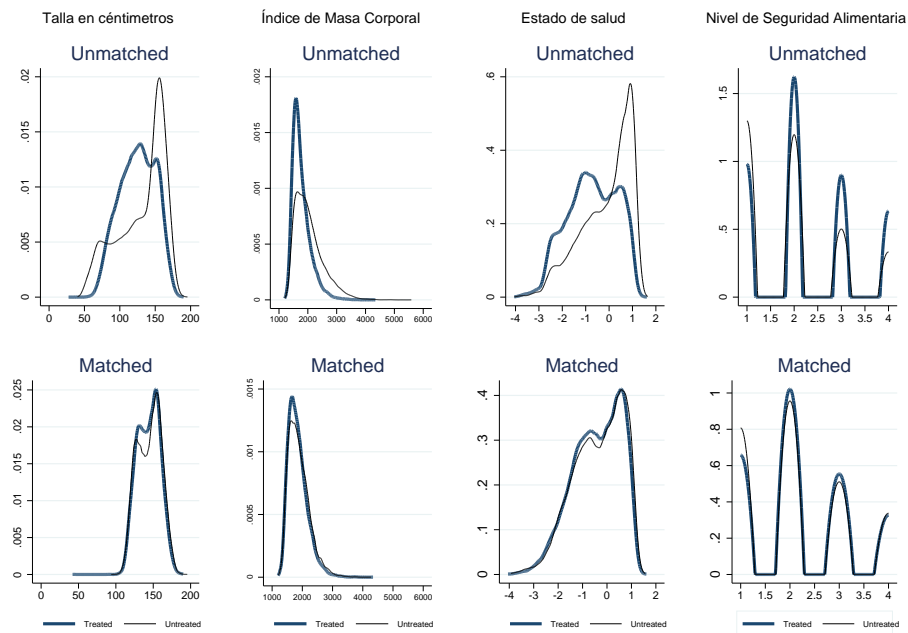
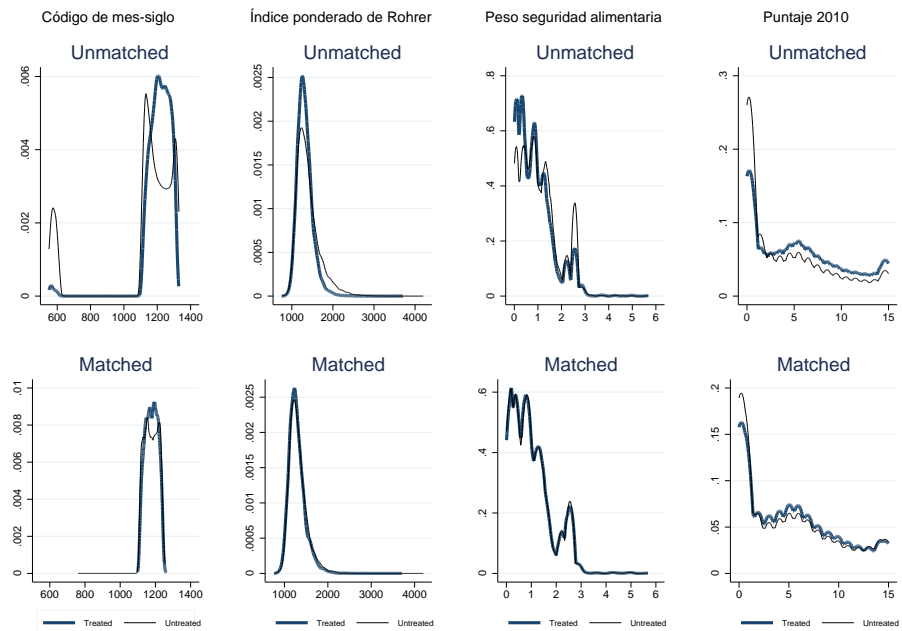


Figura 8: Distribución antes y después del emparejamiento



Cuadro 12: Diferencia de medias entre tratados y controles para variables no incluidas en el PSM

Variable	Sin matching/Con Matching	Tratados	Controles	% Sesgo	P_value
Talla	U	135.31	146.92	-63.5	0.000
	M	144.88	144.59	-3.9	0.101
Peso de Seguridad Alimentaria	U	0.87566	1.0166	-18.6	0.000
	M	0.96969	0.96746	0.3	0.822
Escala de Seguridad Alimentaria	U	2.3057	1.9871	32.3	0.000
	M	2.2191	2.2199	8	0.501
Índice de Rohrer	U	1317.5	1390.3	-30.4	0.000
	M	1289.1	1289.8	-7.8	0.101
Código de mes de nacimiento	U	1189.6	1077	65.7	0.000
	M	1178.7	1177.8	0.5	0.408
Índice de masa corporal	U	1774.7	2041.7	-64.6	0.000
	M	1888	1888.9	-8.9	0.591
Puntaje de S.A	U	5.5035	4.0281	31.1	0.000
	M	5.0581	5.0447	7.2	0.096

Cuadro 13: Prueba de Falsificación

Niño con discapacidad		
Variable	Coef	D. Est
Emparejamiento simple	0.081	0.095
Emparejamiento con reducción de sesgo m=4	0.075	0.098



Cuadro 14: Test de kolmogorov-smirnov para diferencias en distribución. *Modelo de emparejamiento*

Variablen	Sin Matching (U) /Con Matching(M)	P-Value
Edad	U	0.002
	M	0.805
Hombre	U	0.005
	M	0.846
Último grado	U	0.023
	M	0.228
Madre en el Hogar	U	0.000
	M	0.743
Padre en el Hogar	U	0.000
	M	0.522
Educación de la Madre	U	0.027
	M	0.773
Tamaño del hogar	U	0.002
	M	0.107
Vivienda Propia	U	0.004
	M	0.921
Estrato	U	0.000
	M	0.118
Régimen contributivo	U	0.000
	M	0.144
Adulto Mayor en el hogar	U	0.027
	M	0.470
Área Urbana	U	0.000
	M	0.971
Sisbén	U	0.000
	M	0.701
Familias en Acción	U	0.050
	M	0.601
Percibe Ingreso	U	0.081
	M	0.881
Madre con primaria	U	0.002
	M	0.347
Madre con secundaria	U	0.002
	M	0.793
Madre con universidad	U	0.001
	M	0.118

La tabla muestra los p-values del test de kolmogorov-smirnov, donde la hipótesis nula es que las distribuciones sean iguales.

Cuadro 15: Test de kolmogorov-smirnov para diferencias en distribución. *Variables no incluídas en el Propensity Score*

Variablen	Sin Matching (U) /Con Matching(M)	P-Value
Talla	U	0.008
	M	0.247
Peso de seguridad alimentaria	U	0.036
	M	0.708
Escala de seguridad alimentaria	U	0.002
	M	0.2886
Índice de Rohrer	U	0.002
	M	0.613
Código de mes de nacimiento	U	0.008
	M	0.239
Índice de Masa Corporal	U	0.075
	M	0.193
Puntaje de S.A	U	0.007
	M	0.083

La tabla muestra los p-values del test de kolmogorov-smirnov, donde la hipótesis nula es que las distribuciones sean iguales.

Cuadro 16: Test de kolmogorov-smirnov para diferencias en distribución. *Modelo de emparejamiento con corrección de sesgo*

Variable	m=4	m=16	m=64	m=256
Grado Escolar	0.383	0.010	0.000	0.001
Edad_area	0.587	0.101	0.002	0.000
Area_sexo	0.967	0.901	0.026	0.005
Madre en el Hogar	0.965	0.030	0.010	0.001
Padre en el Hogar	0.302	0.007	0.006	0.056
Educación de la madre	0.705	0.176	0.018	0.023
Tamaño del Hogar	0.107	0.031	0.003	0.016
Adulto Mayor en el hogar	0.230	0.022	0.007	0.040
Vivienda Propia	0.488	0.005	0.014	0.000
Estrato	0.750	0.004	0.022	0.000
Régimen Contributivo	0.101	0.009	0.009	0.072
Edad	0.142	0.019	0.002	0.001
Edad2	0.165	0.011	0.018	0.005
Edad3	0.945	0.009	0.015	0.001
Tamaño Hogar3	0.936	0.016	0.007	0.000
Estrato2	0.775	0.022	0.022	0.000
Estrato3	0.101	0.016	0.007	0.000
Area	0.997	0.004	0.007	0.001
Agua	0.555	0.501	0.022	0.001
Etnia	0.270	0.101	0.000	0.000
Sexo	0.997	0.891	0.101	0.000

La tabla muestra los p-values del test de kolmogorov-smirnov para cada diferente número de vecinos m, donde la hipótesis nula es que las distribuciones sean iguales.

Cuadro 17: Test de kolmogorov-smirnov para diferencias en distribución. *Emparejamiento genético*

Variablen	Sin Matching (U) /Con Matching(M)	P-Value
Edad	U	0.002
	M	0.905
Hombre	U	0.005
	M	0.849
Último grado	U	0.023
	M	0.339
Madre en el Hogar	U	0.000
	M	0.889
Padre en el Hogar	U	0.000
	M	0.622
Tamaño del hogar	U	0.002
	M	0.555
Vivienda Propia	U	0.004
	M	0.967
Estrato	U	0.000
	M	0.225
Régimen contributivo	U	0.000
	M	0.453
Adulto Mayor en el hogar	U	0.027
	M	0.667
Área Urbana	U	0.010
	M	0.991
Sisbén	U	0.000
	M	0.101
Familias en Acción	U	0.050
	M	0.336
Percibe Ingreso	U	0.081
	M	0.333
Madre con primaria	U	0.002
	M	0.401
Madre con secundaria	U	0.002
	M	0.661
Madre con universidad	U	0.001
	M	0.166

La tabla muestra los p-values del test de kolmogorov-smirnov, donde la hipótesis nula es que las distribuciones sean iguales.

## 9. Referencias

Akabayashi, H., y Psacharopoulos, G. (1999). The trade off between child labour and human capital formation: A Tanzanian case study. *The Journal of Development Studies*, 35(5), 120-140.

Abadie, A., Drukker, D., Herr, J. L., y Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *Stata journal*, 4, 290-311.

Abadie, A. (2005), “Semiparametric Difference-in-Differences Estimators,” *Review of Economic Studies*, 72, 1–19. [1]

Abadie, A., y Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*.

Acosta. (2013). Evaluación de Operaciones y Resultados para Determinar el Grado de Efectividad del Programa de Alimentación Escolar – PAE. CINDE g-exponencial. Consultores.

Afridi, F. (2010). Child welfare programs and child nutrition: Evidence from a mandated school meal program in India. *Journal of Development Economics*, 92(2), 152-165.

Ahmed, A. U., y Arends-Kuenning, M. (2006). Do crowded classrooms crowd out learning? Evidence from the food for education program in Bangladesh. *World development*, 34(4), 665-684.

Akabayashi, H., & Psacharopoulos, G. (1999). The trade off between child labour and human capital formation A Tanzanian case study. *The Journal of Development Studies*, 35(5), 120-140.

Alderman, Harold., y Bundy, Donald. 2011. School feeding programs and development: are we framing the question correctly?. *The World Bank Research Observer*, lkr005.

Adelman, Sarah, Daniel O. Gilligan, and Kim Lehrer. 2008. How Effective are Food for Education Programs? A Critical Assessment of the Evidence from Developing Countries. IFPRI Food Policy Review 9. Washington, DC: International Food Policy Research Institute.

Álvarez MC, López A, Monsalve J, Giraldo N, Zapata O, Vélez O, et al. Seguridad alimentaria de los hogares de los niños que participan en el programa de Complementación Alimentaria alianza MANA - ICBF. En: Contexto sociodemográfico, estado nutricional y de salud e ingesta dietética

de los niños que participan del programa de Complementación Alimentaria alianza MANA ICBF. Medellín; Divergráficas S.A, 2007. p. 32-131.

Attanasio, O., y Gómez, L. C. (2004). Evaluación Del Impacto Del Programa Familias En Acción-Subsidios Condicionados de La Red de Apoyo Social. Informe Del Primer Seguimiento (Ajustado).

Baland, J. M., y Robinson, J. A. (2000). Is child labor inefficient?. *Journal of Political Economy*, 108(4), 663-679.

Basu, K., y Van, P. H. (1998). The economics of child labor. *American Economic Review*, 412-427.

Basu, K., y Tzannatos, Z. (2003). The Global Child Labor Problem: What do we know and what can we do?. *The World Bank Economic Review*, 17(2), 147-173.

Becker, G. S. (1981). *A Treatise on the Family*. Harvard University Press.

Becker, G. S. (2009). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Beegle, K., De Weerd, J., y Dercon, S. (2006). Poverty and wealth dynamics in Tanzania: Evidence from a tracking survey. World Bank. Washington, DC. Processed.

Behrman, J., Hoddinott, J., Maluccio, J., Quisumbing, A., Martorell, R., y Stein, A. (2003). The impact of experimental nutrition interventions on education into adulthood in rural Guatemala: a longitudinal preliminary analysis. Mimeo

Bernal, R. y Peña, X. (2011). Guía práctica para la evaluación de impacto (No. 361.25 B517). Universidad de Los Andes, Bogotá (Colombia).

Bernal, R., y Fernández, C. (2013). Subsidized childcare and child development in Colombia: Effects of Hogares Comunitarios de Bienestar as a function of timing and length of exposure. *Social Science & Medicine*, 97, 241-249.

Bhalotra, S., y Heady, C. (2003). Child farm labor: The wealth paradox. *The World Bank Economic Review*, 17(2), 197-227.

Blundell, R., y Pistaferri, L. (2003). Income volatility and household consumption: The impact of food assistance programs. *Journal of Human resources*, 1032-1050.

Bundy, Donald, S. Shaeffer, M. Jukes, K. Beegle, A. Gillespie, L. Drake, F.L. Seung-hee, A-M. Hoffman, J. Jones, A. Mitchell, C. Wright, D. Camara, C. Golmar, L. Savioli, T. Takeuchi, and M. Sembene. 2006. "School Based Health and Nutrition Programs." In D. Jamison, J.G. Breman, A.R. Measham, G. Alleyne, M. Claeson, D. Evans, P. Jha, A. Mills, and P Musgrove., eds., *Disease Control Priorities in Developing Countries*. 2nd edn. New York: World Bank and Oxford University Press: 1091–108.

Camacho, A. (2008). Stress and Birth Outcomes: Evidence from Terrorist Attacks in. In Colombia", *American Economic Review*.

Cameron, S. V., & Heckman, J. J. (1998). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts (No. w6385). National Bureau of Economic Research.

Canagarajah, S., y Coulombe, H. (1997). Child labor and schooling in Ghana. World Bank Policy Research Working Paper, (1844).

Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives*, 15(2), 213-238.

Cueto, S., Chinen, M. (2001). Impacto educativo de un programa de desayunos escolares en escuelas rurales del Perú.

Dammert, A. (2005). Does child labor decline with household income? A non-parametric approach. Syracuse University.

Das, J., Do, Q. T., y Özler, B. (2005). Reassessing conditional cash transfer programs. *The World Bank Research Observer*, 20(1), 57-80.

Dehejia, Rajeev (2005). "Practical Propensity Score Matching: A Reply To Smith and Todd." *Journal of Econometrics*, 125, 355–364.

Durrant, J. E. (2011). *Positive Discipline in Every Day Parenting*. Save the Children Sweden and Global Initiative to End all Corporal Punishment of Children. Tercera Edición.

Ersado, L. (2005). Child Labor and Schooling Decisions in Urban and Rural Areas: Comparative Evidence from Nepal, Peru, and Zimbabwe. *World Development*,

Feldman, R. S. (2011). *Child Development*. Pearson Education, Inc. Sixth Edition.

Gertler, P., Patrinos, H. A., y Rodríguez-Oreggia, E. 2012. Parental Empowerment in Mexico: Randomized Experiment of the Apoyos a la Gestión Escolar (AGE) in Rural Primary Schools in Mexico. Washington, DC: The World Bank.

Guarcello, L., Mealli, F., y Rosati, F. C. (2010). Household vulnerability and child labor: the effect of shocks, credit rationing, and insurance. *Journal of population economics*, 23(1), 169-198.

Frankenburg, W.k, Dodds, J., Archer, P., Shapiro, H. y Bresnick, B. 1992. The Denver II: A major revision and restandardization of the Denver Developmental Screening Test. *Pediatrics*, 89, 91-97.

González, L., Claros, S., Cano, G., y Flechas, H. (2009). Comparación evaluación costo beneficio programas nutricionales en Colombia: Familias en Acción y Hogares Comunitarios. Bogotá: Departamento Nacional de Planeación.

Grantham-McGregor, S. M., Walker, S. P., Chang, S. M., y Powell, C. A. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *American Journal of Clinical Nutrition* 66, 247-253.

Grantham-McGregor, S. M., Cheung, Y.B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., y the International Child Development Steering Group. 2007. Developmental Potential in the First 5 Years for Children in Developing Countries. *The Lancet*, Vol 369, 69-70.

Hamelin AM, Habitch JP, Beaudry M. Food insecurity: consequences for the household and broader social implications. *J Nutr*. 1999; 129(2S Suppl):5255-85

Hackett M, Melgar-Quinonez H, Taylor CA, Álvarez Uribe MC. Factors associated with household food security of participants of the MANA food supplement program in Colombia. *Arch Latinoam Nutr*. 2010; 60(1):42-7.

He, H. (2016). Child labour and academic achievement: Evidence from Gansu Province in China. *China Economic Review*, 38, 130-150.



Heady, C. (2003). The effect of child labor on learning achievement. *World Development*, 31(2), 385-398.

Instituto Colombiano de Bienestar Infantil. (2007). Lineamientos Técnicos Programa de Alimentación Escolar.

Islam, Mahnaz, and John Hoddinott. 2009. "Evidence of Intrahousehold Flypaper Effects from a Nutrition Intervention in Rural Guatemala." *Economic Development and Cultural Change* 57(2): 215–38.

Jacoby, Hanan G. 2002. "Is there an Intrahousehold 'Flypaper Effect'? Evidence from a School Feeding Programme." *Economic Journal* 112(476): 196–221.

Jacoby, H. G., y Skoufias, E. (1997). Risk, financial markets, and human capital in a developing country. *The Review of Economic Studies*, 64(3), 311-335.

Johnson, W. R., y Neal, D. (1998). Basic skills and the black-white earnings gap. *The Black-White test score gap*, 480-497.

Jukes, M.C.H., L.J. Drake, and D.A.P. Bundy. 2007. *School Health, Nutrition, and Education for All: Leveling the Playing Field*. Wallingford, Oxfordshire, UK: CAB International.

Kail, R.V. 2004. Cognitive development includes global and domain-specific processes. *Merrill-Palmer Quarterly*, 50, 445-455.

Kazianga, Harounan, Damien de Walque, and Harold Alderman. 2009. "Educational and Health Impact of Two School Feeding Schemes: Evidence from a Randomized Trial in Rural Burkina Faso." *World Bank Policy Research Working Paper* 4976.

Keane, M. P., & Wolpin, K. I. (1997). The career decisions of young men. *Journal of political Economy*, 105(3), 473-522.

Kristjansson, E.A., V. Robinson, M. Petticrew, B. MacDonald, J. Krasevec, L. Janzen, T. Greenhalgh, G. Wells, J. MacGowan, A. Farmer, B.J. Shea, A. Mayhew, and P. Tugwell. 2007. "School Feeding for Improving the Physical and Psychosocial Health of Disadvantaged Elementary School Children." *Cochrane Database of Systematic Reviews Issue 1*, Art. CD004676. ([www.cochrane.org/reviews/en/ab004676.html](http://www.cochrane.org/reviews/en/ab004676.html)).

Kruger, D. (2004). Child labor and schooling during a Coffee Sector Boom: Nicaragua 1993-1998. *Trabajo Infantil: Teoría y Evidencia desde Latinoamérica*.

Labenne, S. (1997). The determinants of child labor in India. Manuscript. Namur, Belgium: Univ. Namur, Center Res. Econ. Development.

Levison, D., Anker, R., Ashraf, S., y Barge, S. (1996). Is child labor really necessary in India's carpet industry?. [Unpublished] 1994. Presented at the Annual Meeting of the Population Association of America Miami Florida May 5-7 1994..

McEwan, P. J. (2013). The impact of Chile's school feeding program on education outcomes. *Economics of Education Review*, 32, 122-139.

Melgar-Quinonez H, Zubieta A, Mk-Nelly B, Nteziyaremye A, Gerardo M, Dunford C. Household Food insecurity and daily per capita food expenditure in Bolivia, Burkina Faso and the Philippines. *Journal of Nutrition*. 2006; 136: 1431S-7S.

Mhurchu, C. N., Turley, M., Gorton, D., Jiang, Y., Michie, J., Maddison, R., & Hattie, J. (2010). Effects of a free school breakfast programme on school attendance, achievement, psychosocial function, and nutrition: a stepped wedge cluster randomised trial. *BMC Public Health*, 10738-743. doi:10.1186/1471-2458-10-738

Murnane, R. J., Willett, J. B., & Levy, F. (1995). The Growing Importance of Cognitive Skills in Wage Determination. *The Review of Economics and Statistics*, (2). 251.

Organización Internacional del trabajo. (2012). *Desarrollo de la Política Pública para la prevención y erradicación del trabajo infantil en Colombia*.

Piaget, J. 1962. *Play, dreams and imitation in childhood*. New York: Norton.

Piaget, J. 1983. Piaget's theory. In W. Kessen (Ed.) P.H. Mussen (Series Ed.) *Handbook of child psychology: Vol 1. History, theory and methods* (pp. 103-128). New York: Wiley.

Powell, C. A., Walker, S. P., Chang, S. M., y Grantham-McGregor, S. M. (1998). Nutrition and education: a randomized trial of the effects of breakfast in rural primary school children. *The American journal of clinical nutrition*, 68(4), 873-879.

Quiroga, B. M. (2006). Trabajo infantil en los niños y jóvenes beneficiarios del programa Familias en Acción: una evaluación de impacto. Fedesarrollo.

Ravallion, M., y Wodon, Q. (2000). Does child labour displace schooling? Evidence on behavioural responses to an enrollment subsidy. *The Economic Journal*, 110(462), 158-175.

Ray, R. (2000). Child labor, child schooling, and their interaction with adult labor: Empirical evidence for Peru and Pakistan. *The World Bank Economic Review*, 14(2), 347-367.

Rodriguez, C., y Sanchez, F. (2012). Armed conflict exposure, human capital investments, and child labor: Evidence from Colombia. *Defence and peace economics*, 23(2), 161-184.

Rosenbaum, P and Rubin, D. (1983). The Central Role of Propensity Score in Observational Studies of Causal Effect. *Biometrika*.

Rosenzweig, M. R., y Wolpin, K. I. (1985). Specific experience, household structure, and intergenerational transfers: farm family land and labor arrangements in developing countries. *The Quarterly Journal of Economics*, 961-987.

Sekhon, J. S., y Diamond, A. (2005). Genetic matching for estimating causal effects. In Unpublished Manuscript. Presented at the Annual Meeting of the Political Methodology, Tallahassee, FL.

Sekhon, J. S., y Mebane, W. R. (1998). Genetic optimization using derivatives. *Political Analysis*, 7(1), 187-210.

Smith, J. A., y Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators?. *Journal of econometrics*, 125(1), 305-353.

Symons, C. W., Cinelli, B., James, T. C., y Groff, P. (1997). Bridging student health risks and academic achievement through comprehensive school health programs. *Journal of school Health*, 67(6), 220-227.

Thomas, D., Beegle, K., Frankenberg, E., Sikoki, B., Strauss, J., y Teruel, G. (2004). Education in a Crisis. *Journal of Development Economics*, 74(1), 53-85.

van Stuijvenberg, M.E., J.D. Kvalsvig, M. Faber, M. Kruger, D.G. Kenoyer, and A.J. Benade.

1999. “Effect of Iron-, Iodine-, and b-carotene-fortified Biscuits on the Micronutrient Status of Primary School Children: A Randomized Controlled Trial.” American Journal of Clinical Nutrition 69: 497–503.

Wahba, J. (2000). Child labor and poverty transmission: no room for dreams. Economic Research Forum for the Arab countries, Iran y Turkey.

Whaley, Shannon, Marian Sigman, Charlotte Neumann, Nimrod Guthrie, Robert E. Weiss, Susan AlberSuzanne P. Murphy. 2003. “The Impact of Dietary Intervention on the Cognitive Development of Kenyan School Children.” Journal of Nutrition 133(11): 3965S–3971S.

Walker, W.A. y Humphries, C. 2005. The Harvard Medical School Guide to Healthy Eating during Prenancy. New York: McGraw-Hill.

Webb P, Coates J, Frongillo EA, Rogers BL, Swindale A, Bilinsky P. Measuring household food insecurity: why it’s so important ant yet so difficult to do. J Nutr. 2006; 136(5): 1404S-8S.

## **10. Apéndice 1: Supuestos del emparejamiento con reducción de sesgo y otros aspectos**

A continuación se detallan los supuestos considerados por Abadie e Imbens (2011), que logran eliminar la correlación entre  $D_i$  y  $Y_i$  y que son útiles para la identificación y estimación del efecto promedio del tratamiento.

*Supuesto S.1.* Sea  $X$  un vector aleatorio de dimensión  $k$  de covariados continuos distribuido en  $\mathbb{R}^k$  con soporte convexo y compacto en  $\mathbb{X}$ , con densidad acotada y acotada a distancia de cero en su soporte.

Este supuesto requiere que todas las variables en  $X$  tengan una distribución continua, sin embargo covariados discretos con un número finito de puntos de soporte pueden considerarse en este análisis condicionando en sus valores. Dicha funcionalidad permite incluir en el análisis covariados discretos, la misma que han aprovechado Mackenzie et al (2010) y los mismos Abadie e Imbens (2006 y 2011).

*Supuesto S.2.* Para casi todo  $x \in \mathbb{X}$ ,

(i) (unconfoundedness)  $W$  es independiente de  $(Y_i(0), Y_i(1))$  condicional en  $X_i = x$ ;

(ii) (Overlap)  $\eta | < Pr(D_i = 1 | X_i = x) < 1 - \eta$ , para algún  $\eta > 0$

Por su parte, S.2(i) establece que condicional en  $X_i$  el tratamiento  $W_i$  es tan bueno como si fuera aleatorio, es decir,  $D_i$  es independiente de los resultados potenciales,  $Y_i(1)$  y  $Y_i(0)$ . Este será el caso si todos los potenciales cofactores son incluidos en  $X$ . Por lo tanto, condicional en  $X_i = x$ , una simple comparación del promedio de resultados es igual al efecto promedio del tratamiento dado  $X_i = x$ . Mientras que S.2(ii) es el supuesto usual de soporte común. La combinación de S.2(i) y S.2(ii) se conoce como ignorabilidad fuerte.

*Supuesto S.3.*  $(Y_i, D_i, X_i)_{i=1}^N$  son series independientes de la distribución de  $(Y, D, X)$ .

A su vez, este supuesto se refiere al proceso muestral.

*Supuesto S.4.* Sea  $\mu_D(x) = \mathbb{E}[Y_i(D) | X_i = x]$  y  $\sigma_D^2(x) = \mathbb{E}[(Y_i - \mu_D(x))^2 | X_i = x]$ . Entonces, (i)  $\mu_D(x)$  son Lipschitz continuos en  $\mathbb{X}$  para  $D = [0, 1]$ , (ii)  $\mathbb{E}[(Y_i(D))^4 | X_i = x] \leq C$  para algún  $C$  finito, para casi todo  $x \in \mathbb{X}$  y (iii)  $\mu_D^2(x)$  está acotado a distancia de cero.

Finalmente, este supuesto recoge condiciones regulares.

Ahora bien, en esta ocasión se considera un modelo de emparejamiento con reemplazo, permitiendo a cada unidad ser usada como una pareja más de una vez. En ese sentido, se emparejan las observaciones haciendo uso de la distancia de Mahalanobis, es decir, si  $\mathcal{L}_m(i)$  es la  $m$ -ésima unidad más cercana a la unidad  $i$  en términos de covariados entonces minimiza la siguiente distancia.

$$\sum_{j: D_j = 1 - D_i} \mathbf{1} \|X_j - X_i\|_A \leq \|X_{\mathcal{L}_m(i)} - X_i\|_A = m \quad (14)$$

Donde  $\mathbf{1}$  es la función indicadora, igual a 1 si la expresión es verdadera y cero en otro caso. Una vez las parejas son encontradas se procede a hacer la estimación. Por otro lado  $A$  es una matriz  $k \times k$  que es escogida para ser la inversa de la matriz de varianzas y covarianzas muestral de los covariados, correspondiente a la métrica de Mahalanobis, es decir:

$$A_{maha} = \left( \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}).(X_i - \bar{X})' \right) \quad (15)$$

Donde  $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$  y  $m$  es la pareja a la que le corresponde ese vecino.

Sea  $\mathcal{J}_M(i) = \mathcal{L}_1(i), \dots, \mathcal{L}_M(i)$  que denota el conjunto de índices para la primera pareja de  $M$  para la unidad  $i$ , para  $M$  tal que  $M \leq N_0$  y  $M \leq N_1$ , es decir, el conjunto de parejas  $M$  es menor al grupo de tratados y controles. Finalmente, sea  $K_M(i)$  el número de veces que la unidad  $i$  es usada como una pareja si hacemos coincidir cada unidad a la pareja de  $M$  más cercana, y

$$K_M(i) = \sum_{l=1}^N \mathbf{1}_{i \in \mathcal{J}_M(l)}$$

Bajo emparejamiento sin reemplazo,  $K_m(i) \in 0, 1$ , pero en una configuración de emparejamiento con reemplazo,  $K_m(i)$  puede tomar un valor entero mayor a 1 si la unidad  $i$  es la pareja más cercana para varias unidades.

El método para encontrar los estimadores bajo esta estrategia consiste en la imputación del resultado potencial faltante (Imbens y Abadie, 2011). Entonces, para cada  $i$  es observable  $y_i(0)$  si  $D_i=0$  ó  $y_i(1)$  en otro caso. El resultado potencial faltante para  $i$  es imputado usando el promedio de los resultados del individuo con el que son emparejados. Lo anterior lleva a que:

$$\hat{Y}_i(0) = \begin{cases} Y_i & \text{si } D_i = 0 \\ \left( \frac{1}{M} \right) \sum_{j \in M(i)} Y_j & \text{si } D_i = 1 \end{cases} \quad (16)$$

y

$$\hat{Y}_i(1) = \begin{cases} \left( \frac{1}{M} \right) \sum_{j \in M(i)} Y_j & \text{si } D_i = 0 \\ (Y_i) & \text{si } D_i = 1 \end{cases} \quad (17)$$

Tal que  $M \leq N_0$  y  $M \leq N_1$ , es decir, el conjunto de parejas  $M$  es menor al grupo de tratados y controles

Usando esta notación, es posible escribir los estimadores del emparejamiento para el efecto

promedio del tratamiento  $\tau$  basado en  $M$  parejas por unidad  $m$ , con reemplazo:

$$\tau_M^m = \left(\frac{1}{N}\right) \sum_{i=1}^N (\hat{Y}_i(1) - \hat{Y}_i(0)) \quad (18)$$

Sin embargo, antes de introducir los estimadores del emparejamiento es importante considerar los de regresión. Sea  $\hat{\mu}_D(x)$  un estimador consistente de  $\mu_D(x)$ . A su vez, un estimador de una regresión de imputación usa  $\hat{\mu}_0(X_i)$  y  $\hat{\mu}_1(X_i)$  para imputar el valor faltante de  $Y_i(0)$  y  $Y_i(1)$  respectivamente. Esto es,

$$\bar{Y}_i(0) = \begin{cases} Y_i & \text{si } D_i = 0 \\ (\hat{\mu}_0(X_i)) & \text{si } D_i = 1 \end{cases} \quad (19)$$

y

$$\bar{Y}_i(1) = \begin{cases} \hat{\mu}_1(X_i) & \text{si } D_i = 0 \\ (Y_i) & \text{si } D_i = 1 \end{cases} \quad (20)$$

Luego, el estimador de la regresión de imputación de  $\tau$  es

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^N (\bar{Y}_i(1) - \bar{Y}_i(0)) \quad (21)$$

A su vez, los autores ratifican lo considerado en Abadie e Imbens (2006) y clasifican como estimadores de una regresión de imputación aquellos para los cuales  $\hat{\mu}_D(x)$  es un estimador consistente de  $\mu_D(x)$ .

Para ilustrar la diferencia entre el estimador simple del emparejamiento, el de la regresión y el el estimador del emparejamiento corregido por sesgo se puede considerar la unidad  $i$  con  $D_i=0$ . Para esta unidad,  $Y_i(0)$  es conocido y  $Y_i(1)$  necesita ser imputado. El estimador simple del matching imputa el outcome potencial faltante  $Y_i(1)$  como:

$$\hat{Y}_i(1) = \frac{1}{M} \sum_{j \in \mathcal{J}_M(i)} Y_j(1) \quad (22)$$

Donde  $\mathcal{J}_M(i) = \mathcal{L}_1(i), \dots, \mathcal{L}_M(i)$  denota el conjunto de índices para la primera pareja de M para la unidad  $i$ . Por otro lado, el estimador imputado por la regresión es este:

$$\bar{Y}_i(1) = \hat{\mu}_1(X_i) \quad (23)$$

## 11. Apéndice 2: Soporte común para las muestras empleadas en los efectos heterogéneos

Figura 9: Propensity Score trabajo Infantil para Hombres

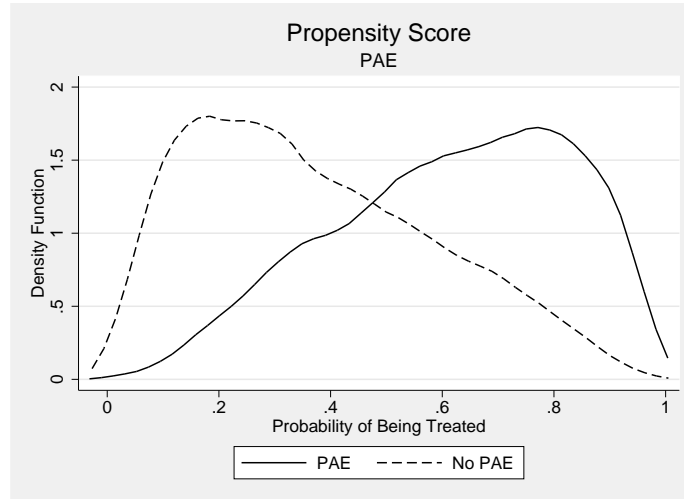




Figura 10: Propensity Score trabajo Infantil para Área Rural

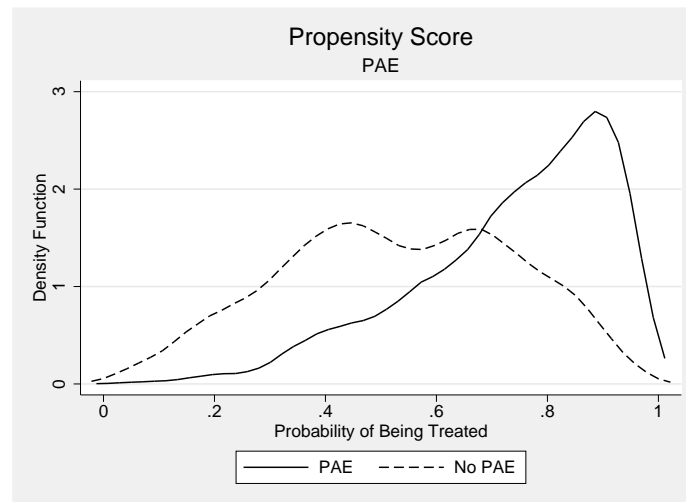


Figura 11: Propensity Score trabajo Infantil para Secundaria

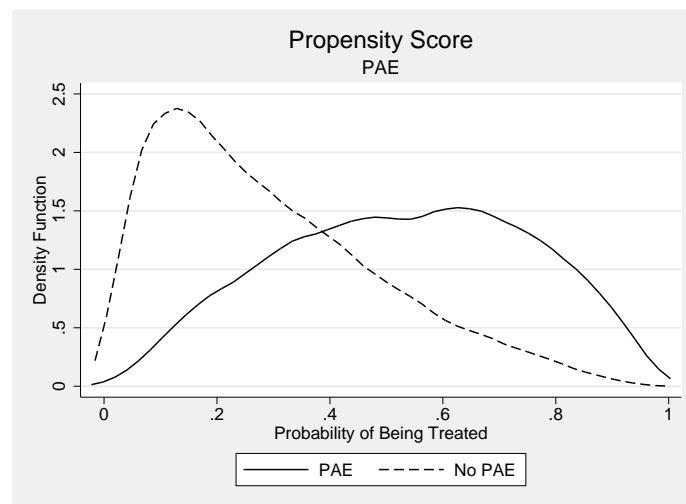


Figura 12: Propensity Score trabajo Infantil para Secundaria y Rural

